

ENCODING OF SOUND-SOURCE ELEVATION BY THE SPIKE PATTERNS OF
CORTICAL NEURONS

By

LI XU

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1999

ACKNOWLEDGMENTS

First of all, I thank my mentor and role model, Dr. John Middlebrooks, for his teaching, guidance, support, and encouragement during my graduate training. The knowledge and experience that I have gained in his laboratory have contributed greatly to the development of my academic career.

I thank the members of my supervisory committee — Drs. Roger Reep, Charles Vierck, Jr., and Robert Sorkin — for their constructive comments as well as critical questions. I thank Dr. David Green who, although retired from the supervisory committee, has provided me continuous help.

I am grateful to have worked with several postdoctoral fellows in Dr. Middlebrooks's laboratory — Drs. Ann Clock Eddins, Shigeto Furukawa, and Ewen Macpherson. Ann helped me to fit in the lab. Shigeto has participated in most experiments and has contributed one good idea after another for my data analysis and final discussion. Ewen has made sense to me of the mysteries of psychophysical modeling in spatial hearing. New students in Dr. Middlebrooks's laboratory — Julie Arenberg and Brian Mickey — have brought fresh thoughts to the lab. Many thanks go to Zekiye Onsan, who has provided the ultimate technical assistance in the lab.

I thank my fellow graduate students — Tony Acosta-Rua, Kellye Daniels, Sean Hurley, Alyson Peel, and Jeff Petruska — for their friendship, and I wish them all the best in their careers.

I thank the Department of Neuroscience for allowing me to do my dissertation research away from Florida, and, equally, I thank the Kresge Hearing Research Institute of the University of Michigan for accepting me to complete my research there and for awarding me a one-year traineeship (funded by NIDCD).

Finally, I would like to thank my friends and my family who I always keep in my heart, for their understanding, patience, and faith throughout the years.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS	ii
LIST OF FIGURES	vi
ABSTRACT	viii
 CHAPTERS	
1 INTRODUCTION	1
2 BACKGROUND.....	4
Acoustical Cues for Sound Localization.....	4
Auditory Cortex: Structure and Function	8
Area A1	8
Area A2	14
AAF.....	15
Area AES	17
Neural Codes for Sensory Stimuli	20
Spike Rate as Neural Codes	20
Spike Timing as Neural Codes	22
3 SENSITIVITY TO SOUND-SOURCE ELEVATION IN NONTONOTOPIC AUDITORY CORTEX.....	28
Introduction.....	28
Methods	30
Results.....	33
General Properties of Sound-Source Elevation Sensitivity	33
Neural Network Classification of Spike Patterns	38
Comparison of Elevation Coding in Areas AES and A2.....	47
Contribution of SPL Cues to Elevation Coding	48
Frequency Tuning Properties and Network Performance	54
Relation between Azimuth and Elevation Coding	58
Discussion	60
Acoustical Cues and Localization in Median Plane	60

A2 versus AES: Elevation Sensitivity and Frequency Tuning	
Properties	63
Correlation between Azimuth and Elevation Coding.....	65
Concluding Remarks	66
 4 AUDITORY CORTICAL SENSITIVITY TO VERTICAL SOURCE	
LOCATION: PARALLELS TO HUMAN PSYCHOPHYSICS	68
Introduction.....	68
Methods	71
Experimental Apparatus	71
Multichannel Recording and Spike Sorting.....	72
Stimulus Paradigm and Experimental Procedure	73
Data Analysis.....	76
Results.....	77
General Properties of Neural Responses to Broadband and	
Narrowband Stimuli	78
Network classification of responses to broadband stimulation.....	80
Neural Network Classification of Responses to Narrowband	
Stimulation	82
The Model of Spectral Shape Recognition.....	86
Correspondence of Physiology with Behavioral Simulation.....	92
Neural Responses to Stimuli Containing a Narrowband Notch.....	97
Comparison of Narrowband Noise Results to Highpass Noise Data.....	100
Elevation Sensitivity by Spike Counts.....	108
Discussion	111
Spectral Features and Elevation Coding	112
Influences of Spectral Notches on Elevation Coding.....	116
Elevation Coding by Spike Counts and Spike Timing	117
Concluding Remarks	119
 5 SUMMARY AND CONCLUSIONS.....	121
 REFERENCES	124
 BIOGRAPHICAL SKETCH.....	132

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
3.1. Spike-count-versus-elevation profiles	34
3.2. Distribution of depth of modulation of spike count by elevation	36
3.3. Distribution of the range of elevations over which spike counts greater than half maximum were elicited	37
3.4. Distribution of locations of best-elevation centroids	39
3.5. Raster plot of responses from two AES units (A: 950531 and B: 950754) and an A2 unit (C: 970821)	40
3.6. Network performance of the same unit (950531) as in Figure 3.5A	41
3.7. Network performance of the same unit (950754) as in Figure 3.5B	43
3.8. Network performance of the same unit (970821) as in Figure 3.5C	44
3.9. Distribution of elevation coding performance across the entire sample of units	46
3.10. Comparison of network performance of A2 and AES units	48
3.11. Sound levels and neural network performance	50
3.12. Percentage of unit sample activated as a function of stimulus tonal frequency	55
3.13. Frequency tuning bandwidth and neural network performance	57
3.14. Correlation between network performance in azimuth and elevation	59
4.1. Unit responses elicited by broadband and narrowband noise (unit 9806C02)	79
4.2. Network analysis of spike patterns of the same unit (9806C02) as in Figure 4.1	81
4.3. Unit responses elicited by broadband, narrowband, and notched noise (unit 9806C16)	84
4.4. Network estimates of elevation	85
4.5. Network analysis of spike patterns and model predictions in response to narrowband stimulation	87
4.6. Head-related transfer functions (HRTFs) in the median plane measured from left ears of 3 cats	88
4.7. Spectral differences between the narrowband stimulus spectra and HRTFs	90
4.8. Correspondence between model prediction and network outputs	93
4.9. Distribution of percent correct for all narrowband center frequencies across the sample of units	96
4.10. Network analysis of spike patterns elicited by notched noise	99
4.11. Unit responses elicited by broadband, narrowband, and highpass noise (unit 9811C03)	101
4.12. Comparison of network classification of the spike patterns elicited by	

narrowband and highpass noise	103
4.13. Sum of the squared differences (SSD) of network outputs	105
4.14. Distribution of percentile of matched SSD across the sample of units	107
4.15. Accuracy of elevation coding by spike counts and by full spike patterns	109
4.16. Network classification of spike counts elicited by narrowband sounds	110

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

ENCODING OF SOUND-SOURCE ELEVATION BY THE SPIKE PATTERNS OF
CORTICAL NEURONS

By

Li Xu

May 1999

Chairman: John C. Middlebrooks
Major Department: Neuroscience

Previous studies have demonstrated that the spike patterns of auditory cortical neurons carry information about sound-source location in azimuth. The question arises as to whether those neurons integrate the multiple acoustical cues that signal the location of a sound source, or whether they merely demonstrate sensitivity to a specific parameter that covaries with sound-source azimuth, such as interaural level difference. We addressed that issue by testing the sensitivity of cortical neurons to sound locations in the median vertical plane, where interaural difference cues are negligible. We also tested whether and how cortical neurons use spectral information to derive their elevation sensitivity. The study involved extracellular recording of units in the nontopographic auditory cortex (areas AES and A2) of chloralose-anesthetized cats. Broadband noise and various spectrally-filtered stimuli were presented in an anechoic room from 14 locations in the vertical midline in 20° steps, from 60° below the front horizon, up and

over the head, to 20° below the rear horizon. Artificial neural networks were used to recognize spike patterns, which contain both the number and timing of spikes, and to thereby estimate the locations of sound sources in elevation. The network performance was fairly accurate in classifying spike patterns elicited by broadband noise. Using the same neural network that was trained with spike patterns elicited by broadband noise, we presented spike patterns elicited by spectrally-filtered noise and recorded network estimates of the locations in elevation of those stimuli. This procedure could be considered as the physiological analog of asking a psychophysical listener to report the apparent location of a spectrally-filtered noise. The network elevation estimates based on spike patterns elicited by narrowband and highpass noise exhibited tendencies similar to localization judgments by human listeners. A quantitative model derived from comparison of the stimulus spectrum with the external-ear transfer functions of individual cats could successfully predict the region in elevation that was associated with narrowband noise. These results further support the theory that full spike patterns (including spike counts and spike timing) of cortical neurons code information about sound location and that such neural responses underlie the localization behavior of the animal.

CHAPTER 1 INTRODUCTION

The auditory cortex is essential for sound localization behavior. Human patients with unilateral temporal lobe lesions have difficulties in localizing sounds from the side contralateral to the lesion (Greene 1929; Klinton and Bontecou 1966; Sanchez-Longo and Forster 1958; Wortis and Pfeiffer 1948). Experimental ablations of the cat's auditory cortex also result in deficits in localization of sound sources presented on the side contralateral to the lesion (Jenkins and Masterton 1982). Despite sustained effort in neurophysiological studies of the auditory cortex, the cortical codes for sound localization are still not well understood.

Studies of the optic tectum in the barn owl (Knudsen 1982) and the superior colliculus in mammals (Middlebrooks and Knudsen 1984; Palmer and King 1982) show evidence of single neurons that are selective for sound-source location. The neurons' preferred sound-source locations vary systematically according to the locations of the neurons within the midbrain structure. Therefore, the working hypothesis for most studies of the auditory cortex has been that there exists a topographic code for sound localization in the auditory cortex (Brugge et al. 1994; Clarey et al. 1994; Imig et al. 1990; Middlebrooks Pettigrew 1981; Rajan et al. 1990b). Unfortunately, results reported from the aforementioned studies have not produced evidence to support such a hypothesis.

In 1994, Middlebrooks and colleagues proposed an alternative hypothesis that a distributed code exists for sound localization in the auditory cortex. Studies in his laboratory have shown that spike patterns (spike counts and spike timing) of the auditory cortical neurons carry information about sound-source location (Middlebrooks et al. 1994, 1998; Xu et al. 1998). The essence of the hypothesis of the distributed code for sound localization is that the activity of each individual neuron can carry information about broad ranges of location and that accurate sound localization is derived from information that is distributed across a large population of neurons.

The present study extended that line of research in Middlebrooks's laboratory and expanded the observation from the horizontal plane to the vertical plane. In the central nervous system, the computational processes for sound localization in the vertical plane are different from those involved for sound localization in the horizontal plane, due to different acoustical cues that are used for localization in the two dimensions. Interaural difference cues (i.e., interaural time difference and interaural level difference) are used for horizontal localization, whereas spectral shape cues are used for vertical localization and front/back discrimination. The computational processes for those cues are parallel and segregated as early as in the cochlear nucleus and all the way throughout the brainstem. The present study was designed to address whether the cortical neurons that have previously been shown to code azimuth integrate the multiple acoustical cues that signal the location of a sound source, or whether they merely demonstrate sensitivity to a specific parameter that covaries with sound-source azimuth, such as interaural level difference. Manipulation of source spectra can confound spectral shape cues for vertical localization. Listeners make systematic misjudgments when asked to localize spectrally-

manipulated noise. Since interaural difference cues are still intact, such a spectral manipulation does not cause error in horizontal localization. Thus, manipulation of source spectra provides a way to test more directly that the cortical neurons utilize the spectral shape cues to code sound-source elevation and that their activities are closely related to the localization behavior of the animal. We studied the changes in the elevation sensitivity of the cortical neurons under the conditions of spectrally-manipulated noise stimulation.

The remainder of the document is organized in the following manner. Chapter 2 reviews the acoustical cues for sound localization with an emphasis on the vertical and front/back dimensions. It also provides a background on the structure and function of the auditory cortex followed by a short review on the cortical codes for sensory stimuli with special attention to the coding of stimuli by the timing of spikes. Two subsequent chapters describe two major research projects that deal with elevation coding in the auditory cortex, each with detailed introduction, methods, results, and discussion. Chapter 3 describes the sensitivity to sound-source elevation in the nontopographic auditory cortex. Chapter 4 describes the responses of auditory cortical neurons to spectrally-manipulated noise stimuli that produce localization illusion. Finally, Chapter 5 provides a brief summary and conclusions from the present research.

CHAPTER 2 BACKGROUND

Acoustical Cues for Sound Localization

Unlike visual space that is mapped on the retina in a point-to-point fashion, sound-source locations are not mapped directly onto the ear. Instead, locations must be computed by the brain from sets of acoustical cues that result from the interaction of the incident sound wave with the head and external ears. Azimuth information is derived at high frequencies from the interaural level differences (ILDs) and at low frequencies from interaural phase differences (IPDs). Those binaural difference cues, however, are ambiguous in distinguishing the vertical and front/back locations (i.e., the elevation). In the median sagittal plane, for example, ILD and IPD values are zero at all locations, if the head is perfectly symmetrical. Off the median plane, ILD and IPD are constant for locations that fall on the surface of virtual cones centered on the interaural axis. Thus, Woodworth (1938) coined the term of "cone of confusion." Batteau (1967) was one of the first to draw our attention to the pinna-based spectral cues as a necessary factor to disambiguate the position around the cone. The convoluted surface of the pinna and concha differentially modify the frequency spectrum of the incoming acoustical signal depending on the angle of incidence of the signal. The spectral features, or spectral shape cues, that result from the modification by the pinna, including spectral peaks and notches, vary systematically with sound-source locations (Shaw 1974; Mehrgardt and

Mellert 1977; Humanski and Butler 1988; Middlebrooks et al. 1989; Wightman and Kistler 1989). The frequencies of the spectral peaks and notches increase as sound-source locations are shifted from low to high elevation, both in the front and rear locations. The peaks and notches grow smaller at high elevations (above $\sim 70^\circ$), resulting in a relatively less transformed spectra for sources above the head. There is significant individual variation in the spectral shape cues due to the physical shape and size differences of the pinnae and heads among subjects (Middlebrooks 1999a).

Several lines of evidence from psychophysical studies indicate that spectral shape cues are the major cues for vertical localization. For example, vertical localization is most accurate when the stimulus has a broad bandwidth that contains energy at 4 kHz and above (Butler and Helwig 1983; Gardner and Gardner 1973; Hebrank and Wright 1974b; Makous and Middlebrooks 1990; Roffler and Butler 1968). Spectral shape cues from one ear seem to be sufficient for vertical localization. Vertical localization with a single ear tested by plugging the other ear is almost accurate as with both ears (Hebrank and Wright 1974a; Oldfield and Parker 1986). Patients who have congenital deafness in one ear but normal hearing in the other show accurate vertical localization (Slattery and Middlebrooks 1994). However, a recent virtual localization study revealed some discrepancies in monaural localization between free-field results and virtual-source results (Wightman and Kistler 1997). In that study, vertical localization was eliminated using monaurally-delivered virtual source sounds.

There are numerous studies on how localization is affected by perturbing, obscuring, or removing the spectral shape cues. Gardner and Gardner (1973) measured median plane localization accuracy as listeners' pinnae were gradually occluded with

rubber inserts. Performance was progressively degraded by various degrees of occlusion. These effects were also observed by Fisher and Freedman (1968), who bypassed the listener's pinnae with inserted tubes. A recent study by Hofman and colleagues (1998) offered an intriguing new insight into how the brain learns the transfer functions of the ears. Those researchers modified the subjects' spectral shape cues by reshaping their pinnae with plastic molds. The localization of sound elevation was dramatically degraded immediately after the modification. After six weeks of wearing these molds continuously, though, all subjects seemed to have learned the transfer functions of the new ears, so their vertical localization with the new ears was normal again. More interestingly, learning the new spectral shape cues did not interfere with the neural representation of the original cues, as the subject could localize sounds with both normal and modified pinnae (Hofman et al. 1998).

Bandpassing the acoustic signal is another commonly-used method to either partially or completely remove spectral shape cues from the signal depending on the bandwidth of filter. In the case of tonal stimulation, the source spectrum consists of a single sinusoid component. Roffler and Butler (1968) used tonal signals in their studies of median plane localization. They demonstrated that the apparent elevation of a source depended on its frequency and was independent of its actual position. Some other experiments were performed with narrowband noise stimuli. Blauert (1969/1970) presented 1/3-octave noise from the median plane and showed that the center frequencies of the noise determined whether the apparent position was in front, above or behind. Similar effects were shown by Butler and Helwig (1983) using 1-kHz-wide noise bands with center frequencies ranging from 4 to 14 kHz. A final example of narrowband

localization is described by Middlebrooks (1992). In his experiment, subjects reported a compelling illusion of an auditory image located at an elevation that was determined by the center frequency of the 1/6-octave-wide narrowband sounds, not by the actual source location. A typical subject, for instance, consistently reported an image high and in front when the center frequency was 6 kHz and low and to the rear when the center frequency was 10 kHz. A model that incorporated measurement of the external-ear transfer functions could predict the reported sound locations. In such a model, similarity between the spectra of narrowband stimuli and the external-ear transfer functions was calculated by way of correlation. Localization judgments of the subjects were biased to locations for which the external-ear transfer function most closely resembled the stimulus spectrum (Middlebrooks 1992).

It is worth noting that disruption of spectral shape cues does not affect accurate localization in azimuth (Hofman et al. 1998; Kistler and Wightman 1992; Middlebrooks 1992, 1999b; Oldfield and Parker 1984). It seems that interaural difference cues and spectral shape cues are utilized independently to derive sound-source azimuth and elevation, respectively. The brain is therefore capable of integrating multiple acoustical cues, including ILDs, IPDs, and spectral shape cues, to synthesize the sound locations. How the brain interprets the spectral shape cues is a puzzling question. Models of sound localization support the concept of a central repository of direction templates, derived from the directional transformation of the external ears (Macpherson 1998; Middlebrooks 1992; Zakarauskas and Cynader 1993). In such a theory, the frequency spectrum of an incoming sound is compared to each of the templates, and the one that matches the best then signals the direction of the incoming sound.

Auditory Cortex: Structure and Function

This section describes the morphological organization of the auditory cortex, i.e., the laminar characteristics and the thalamic connections. Focus then moves to the physiological representations in the auditory cortex, including tonotopic arrangement, binaural processing, and sound localization. This review will consider primarily studies in the cat, the species used in the present research.

The cat's auditory cortex is displayed on the lateral surface of the brain. Based on cytoarchitectural characteristics and physiological properties, the auditory cortex is divided into subregions. They are the primary auditory cortex (A1), the second auditory cortex (A2), the anterior auditory field (AAF), the dorsal posterior (DP), posterior (P), ventral posterior (VP), ventral (V), and temporal (T) auditory fields, and the anterior ectosylvian sulcus area (areas AES) (Clarey and Irvine 1986; Imig and Reale 1980). The most complete studies have been done in areas A1, A2, AAF, or AES.

Area A1

The primary auditory cortex is characterized by an overall high packing density in layers II, III and IV of the six layers. The high density of granular cells gives the cortex the term koniocortex, or "dust cortex." The human primary auditory cortex is a 900 - 1600 mm² area of classic koniocortex along the transverse temporal gyri of Heschl, corresponding to area 41 (Brodman 1909). It is surrounded by nonprimary cortex that can be subdivided into four or five areas. In the cat, A1 is located in the dorsal middle ectosylvian gyrus. The distinction of A1 from other auditory cortical areas can be made in sections stained for cell bodies by the light band of the inner sublayer of layer V (Rose

1949). Detailed description of the A1 cytoarchitecture was further provided by Winer (1992). The molecular layer (layer I) is remarkable for its few neurons. The bulk of its connections are with the apical dendrites of deeper-lying neurons or within layer I. The external granule cell layer (layer II) has a wide range of both pyramidal and nonpyramidal neurons, a columnar and vertical organization that is conserved in the deeper layers, and significant neurochemical diversity. Its principal connections are with adjacent nonprimary auditory areas, and it provides local interlaminar projections with layers I-III. The external pyramidal cell layer (layer III) has a complex set of intrinsic and extrinsic connections, including relations with the auditory thalamus and ipsilateral as well as contralateral auditory cortices. This is reflected in its diverse neuronal architecture. The pyramidal cells of various sizes that are more common in the deeper one-half represent the most conspicuous population in this layer. Many commissural cells of origin lie in this layer. The granule cell layer (layer IV), only about 250 μm thick, represents one-eighth of the cortical depth. Its connectivity is dominated by thalamic, corticocortical, and intrinsic input. It also receives projections from the commissural system but does not send fibers to the system like layer III does. The vertical column organization is particularly obvious in this layer. The internal pyramidal cell layer (layer V) is has a cell-sparse, myelin-rich outer half (Va), and an inner half (Vb) with many medium-sized and large pyramidal cells. It is the source of connections to the ipsilateral nonprimary auditory cortex, the contralateral A1, the auditory thalamus and the inferior colliculus. The multiform layer (layer VI) contains the most diverse neuronal population within A1, consisting of at least nine readily recognized types of cells (Winer 1992).

The major thalamic input to A1 comes from the ventral division of the medial geniculate body (MGB). This specific auditory relay system ends predominantly in layer III and IV (Winer 1992). The thalamocortical and corticothalamic A1 projections are highly reciprocal (Andersen et al. 1980). In addition, the connections between MGB and A1 preserve the systematic topography. For example, injection of anterograde tracer into A1 results in a sheetlike labeling in the ventral division of the MGB and the labeling sites change systematically with the central tuning frequencies of the injection sites. A1 also receives minor input from a nontopographic thalamic nucleus (medium-large cell division of the medial division) (Morel and Imig 1987).

The tonotopic organization of A1 in the cat was first demonstrated at the single-cell level by Merzenich and associates (1973, 1975). Frequency is represented across the mediolateral dimension of A1 cortex as isofrequency bands. On an axis perpendicular to this plane of representation, the best frequencies change as a simple function of cortical location. Low frequencies are represented posteriorly, and high frequencies anteriorly. The frequency tuning curves of the vast majority of the A1 neurons are narrow, with the sharpest tuning at higher best frequencies (Phillips and Irvine 1981). Along the isofrequency contour, gradients of tuning sharpness exist. The sharpest frequency tuning is found near the center of the mediolateral extent of A1, and the sharpness of tuning gradually decreases toward the medial and lateral border of A1 as revealed by multiple-unit recordings (Schreiner and Mendelson 1990). In single unit study, the gradient in bandwidth at 40 dB above minimum threshold (BW40) exists in the dorsal half of A1 (A1d), but the ventral half of A1 (A1v) shows no clear BW40 gradient (Schreiner and Sutter 1992). It is a common observation that within the same vertical penetration into

A1, the best frequency is remarkably constant. The cortical area that represents the higher frequencies is disproportionately larger than that represents the lower frequencies, suggesting that more neural machinery of the cat is devoted to encode or extract information relevant to high frequencies.

The representation of a "point" on the sensory epithelia of the cochlea as a "band" of cortex suggests that some other parameter of the auditory stimulus is functionally organized along the isofrequency dimension. There is evidence that groups of neurons with different binaural response properties are segregated with an A1 isofrequency band. More than 90% of the neurons encountered in A1 can be classified into either the excitatory/excitatory (EE) or excitatory/inhibitory (EI) interaction class (Middlebrooks et al. 1980). Typically, a cortical neuron is excited by sound stimulus from the contralateral ear. If stimulus from ipsilateral side excites the neuron and binaural stimulus displays facilitation in the neuronal responses, this neuron is an EE neuron. Otherwise, if ipsilateral stimulation does not excite the neuron and binaural stimulation produces a weaker response, then the neuron is an EI neuron. All neurons encountered along a given radial penetration are of the same binaural response class. In a surface view, neurons of the same binaural response properties aggregate to form patches. Patches formed by the two types of cells are organized in strips running roughly at right angles to the isofrequency contours (Middlebrooks et al. 1980). The thalamic sources of input to these binaural response-specific bands are strictly segregated from each other in the ventral division of the MGB, as identified with retrograde tracers (Middlebrooks and Zook 1983). The functional roles of the binaural topographic organization are unclear.

One hypothesis is that EI regions are responsible for the processing of spatial location information and EE regions for frequency pattern analysis (Middlebrooks et al. 1980).

Early studies by Middlebrooks and Pettigrew (1981) examined the functional organization pertaining to sound localization within A1. Single units were recorded while tonal stimuli were presented in a free sound field. The receptive fields were mapped by plotting boundaries of spatial regions within which stimuli elicited a given neural response. About half of the neurons encountered were location-insensitive or omnidirectional. Two discrete populations of cells could be identified from the pool of the location-selective units. One was hemifield units which responded to sounds presented in the contralateral sound field; the other was axial units which had small, complete circumscribed receptive fields. The axial units had high frequency tuning, and their receptive fields reflected the directionality of the contralateral ear at those frequencies. It is noteworthy that no systematic map of sound space was found in A1 of the cat. Rajan et al. (1990a) found that neurons were sensitive to contra-field, ipsi-field or central-field and neurons of the same type tended to cluster together along the frequency-band strip. However, there were often rapid changes in the azimuth tuning type in units isolated over short distances even though their electrode steps were usually 100 μm and sometimes 50 μm . A1 was found not to be organized in a point-to-point pattern for the sound-source azimuth. Using noise bursts as stimuli, Imig and colleagues (1990) also found that neighboring units exhibited similar azimuth and stimulus level selectivity, suggesting that modular organizations might exist in A1 related to both azimuth and level selectivity. There is a clear relationship between the nonmonotonic rate-level function and the strength of the directionality. That is, virtually all of the cells

in A1 that have the most strongly nonmonotonic level functions are also sensitive to azimuth. Since similar property was not found in the ventral nucleus of the MGB, they concluded that the linkage between azimuth sensitivity and nonmonotonic level tuning emerged in the cortex (Barone et al. 1996).

Recently, a topography of the monotonicity of rate-level functions in cat A1 was revealed (Sutter and Schreiner 1995). The amplitude selectivity varies systematically along the isofrequency contours. Clusters sharply tuned for intensity (i.e., nonmonotonic clusters) are located near the center of the contour. A second nonmonotonic region is several millimeters dorsal to the center. The lowest thresholds of single neurons are consistently located in the nonmonotonic regions. The scatter of single-neuron intensity threshold is smallest at these locations. Although the nonmonotonic neurons have been shown to be predominantly directionally sensitive (Imig et al. 1990), the restricted intensity response and threshold range would not favor them for encoding intensity-independent sound location. However, the response properties of neurons in the dorsal part of A1 are of interest in the context of sound localization. Sutter and Schreiner (1991) recorded single-unit frequency tuning curves in A1. About 20% of the neurons had multi-peaked tuning curves and 90% of them were in the dorsal part of A1. Inhibitory/suppressive bands, as demonstrated with two-tone paradigm, were often present between peaks. It was suggested that these neurons might be sensitive to specific spectrotemporal combinations in the acoustic input and might be involved in complex sound processing. It is an attractive idea that these subpopulations of neurons in the dorsal part of A1 are particularly suitable for detecting the spectral notches that are flanked by two spectral peaks or plateaus. Because spectral notches have been indicated

to be important acoustical cues for localization in elevation, it might be worthwhile to investigate the coding of elevation by these neurons in our future experiments.

Area A2

A2 is located ventral to A1 on the middle ectosylvian gyrus, extending at least 6 mm ventrally from A1. The transition area between A1 and A2 defined physiologically has a width of about 0.5 - 1 mm, concordant with a gradual change of the cytoarchitecture of the border (Schreiner and Cynader 1984). A2 has a distinctive cytoarchitecture arrangement: there are fewer of the pyramidal cells characteristic of layer III in A1, the density of neurons is more or less uniform throughout, except in layer Vb, and large or giant pyramidal neurons mark layer Va. Nevertheless, layer IV is dominated by small, round cells, and the columnar arrangement evident in A1 is conserved here as well (Winer 1992).

A2 loci are thalamocortically and corticothalamically connected with the caudal dorsal nucleus, the ventral lateral nucleus of the ventral division, and the medial division of the MGB. The dorsal division projections are the heaviest of all. These connections are largely segregated from those between A1 and MGB. Injection studies revealed no apparent systematic topography of A2 projection to and from the MGB nuclei. While the connections between A1 or AAF and the ventral division of the MGB is termed the "cochleotopic system," the connections between A2 and the MGB is called the "diffuse system" (Andersen et al. 1980).

A2 neurons are much more broadly tuned in frequency than A1 neurons. There is a gradual transition from sharply tuned A1 neurons to broadly tuned A2 neurons on the border of A1 and A2. Typical A2 neurons are slightly less sensitive to tonal stimuli than

A1 cells and are almost equally sensitive across a broad range of frequencies, commonly spanning several octaves. Therefore, the tonotopic organization within A2 concordant with A1 in orientation is significantly blurred by the strong variability of the characteristic frequencies, isolated low-frequency islands, and increasing bandwidth of the frequency receptive fields (Andersen et al. 1980; Schreiner and Cynader 1984). A2 is bordered posteriorly by tonotopically organized regions of cortex (P and VP) (Andersen et al. 1980).

In terms of binaural interactions, the segregation of EE and EI responses has also been demonstrated in A2, but grouping of "like" responses tends to be highly variable in shape and orientation between animals as compared to A1. The proportion of EO (no interaction, monaural only) neurons in A2 (~24%) is slightly larger than that in A1 (~18%) (Schreiner and Cynader 1984). Discharges of EO neurons are determined by stimulation of one ear (usually contralateral side) and are unaffected by simultaneous stimulation of the other ear. Therefore, their binaural responses are indistinguishable from the monaurally-evoked responses from the sensitive ear.

AAF

AAF is located anterior to A1 on the middle and anterior ectosylvian gyri. In AAF, the neuronal density is somewhat lower than that in A1 and the cells are slightly larger, the pyramidal cell populations in layer IIIa and Va have larger somata than their A1 counterparts, and the cell-poor part of Vb is reduced. In addition, layer IV contains a significant number of pyramidal cells, unlike layer IV in A1 (Winer 1992).

The systematic topography of the thalamocortical and corticothalamic reciprocal projections of AAF with the auditory thalamus are similar to the A1 connections

(Andersen et al. 1980). However, the connections with the ventral division of the MGB are weaker than in A1. The major tonotopic input comes from the lateral part of the posterior group of thalamic nuclei (Po). A2 also receives major input from the nontonotopic thalamic nucleus (medium-large cell region of the medial division) (Morel and Imig 1987).

In AAF, there is a clear tonotopic organization which is a mirror image of that in A1. High frequencies are oriented dorsoventrally along the border with the high-frequency region of A1; lower frequencies are represented in the more rostral cortex. Comparison of the properties of AAF and A1 shows that these two areas are similar in many important features, including unit response properties, short latency, and disproportionally greater representation of higher frequencies. They also share some common thalamocortical inputs. These similarities suggest that AAF is not a "secondary" cortical field, but rather that it and A1 are parallel processors of ascending acoustical information (Knight 1977).

Phillips and Irvine (1982) obtained data on the binaural interactions of 40 AAF neurons. The binaural interactions of AAF neurons were qualitatively similar to those of A1 neurons, but they regarded the data as preliminary due to the small number of neurons studied.

Azimuthal tuning of AAF neurons was measured by Korte and Rauschecker (1993). Spatial tuning of individual neurons as defined by spatial tuning index which was simply the ratio between the minimal and maximal responses from all 7 azimuth locations (-60 to $+60^\circ$ in 20° step) was found not to be different from that of AES neurons. This study was done in only two cats and the number of AAF neurons versus AES neurons

studied was not reported. Certainly, more studies need to be done before any conclusions on the functional organization of AAF in sound localization can be drawn.

Area AES

Area AES is located on the banks and fundus of the anterior ectosylvian sulcus. It is a multiple-modality sensory cortex where neurons responsive to somatosensory, auditory, and visual stimulation are apparently intermingled throughout both banks and fundus of the AES. But it is still controversial whether there are modality-specific (pure visual or pure somatosensory) subregions and the size of those regions within both banks and fundus of AES (see Meredith and Clemo 1989; Clarey and Irvine 1990a). Barbiturate anesthesia, which has been shown to suppress the auditory responses, was considered to be the reason for the discrepancy among different studies (Clarey and Irvine 1990a).

As would be expected for a multisensory cortex, area AES has a wide range of inputs from the thalamus and other cortical regions. Roda and Reinoso-Suarez (1983) studied the thalamic projections to the cortex of AES by the use of retrograde labeling with a direct visual approach to the AES region. It was shown that all labeled neurons in the thalamus were ipsilateral to the injection. The thalamic afferents originated from the ventromedial thalamic nucleus (VM), lateral medial subdivision of the lateral posterior-pulvinar complex (LM), suprageniculate nucleus (Sg), posterior thalamic nuclear group (Po), and magnocellular (or medial) division of the MGB. A small number of labeled neurons was found in the ventral part of the lateral posterior nucleus (LP), VA/VL, MD, and intralaminar nuclei. Slightly different patterns of these thalamocortical connections were observed depending on the portion of the AES region considered. Clarey and

Irvine (1990b) used a physiological guide to inject horseradish peroxidase into the acoustically responsive regions of the AES. The labeling of the medial division of MGB (i.e., the magnocellular division) and other thalamic nuclei were similar to previously described results. The posterior group of thalamic nuclei (Po), a tonotopically organized auditory thalamus, was also found to project to area AES. Since no neurons in area AES were found to show sharp frequency tuning, some degree of convergence of the input from Po must have occurred. No input from the ventral MGB was described.

The cortical input to area AES arises from a number of unimodal and multisensory areas, with a dominant input from the cortex of the suprasylvian sulcus (SSS), which contains several extrastriate visual fields and to a lesser extent some anterior multimodal regions. Area AES also receives input from contralateral AES and contralateral SSS (Clarey and Irvine 1990b; Reinoso-Suarez and Roda 1985). It is not clear whether area AES receives input from other auditory cortex. A recent report did show that AES neurons projected to auditory cortical areas A1 and A2, and temporal (T) auditory field. In the coronal sections of A1, the labeling appeared in patches. When the sections were aligned and serially arranged, the patches formed bands that extended in a rostrocaudal direction across A1 (Miller and Meredith 1998).

Area AES receives input from the motor regions of the thalamus and cortex (Reinoso-Suarez and Roda 1985); therefore, it might be involved in functions that require sensorimotor integration. This speculation was supported by the fact that area AES has dense projection to deep layers of the superior colliculus (SC) (Meredith and Clemo 1989). In the anterograde and retrograde labeling study, Meredith and Clemo (1989) demonstrated that of the auditory cortices (A1; A2; areas A, P, VP, and AES),

only area AES projected to the SC. Auditory SC neurons responded to electric stimulation of the area AES only. However, neither anatomical nor physiological techniques revealed a clear topographic relationship between the area AES and the SC but suggested instead a diffuse and extremely divergent/convergent projection.

No tonotopic organization has been identified in the area AES. The following characteristics of AES cells distinguish them from the bordering A1 and AAF cells: a loss of sharply tuned responses and the appearance of broad or irregular high-frequency tuning, an increase in the latency of response, an increase in the strength of the suprathreshold response to noise, and the advent of response to visual stimulation (Clarey and Irvine 1986, 1990a). The distinction between the AES neurons and A2 neurons is less clear cut. Generally, the AES neurons are more responsive to noise and some are responsive to visual stimulation. When tested for binaural interactions, the AES neurons have predominantly EE responses (Clarey and Irvine 1990a).

Korte and Rauschecker (1993) reported that more than half of the neurons they recorded from the AAF and area AES were "directional." Preliminary data from the same laboratory showed that the neurons' preferred azimuth changed continuously over a certain range, until it jumped discontinuously. A piecewise continuous representation of location preference in the auditory cortex was suggested (Henning et al. 1995). One of the obvious limitations of their work is that azimuth sensitivity was measured within only 60° of the frontal midline. A complete account of the experiment is still not available. Middlebrooks and collaborators (1998) recorded the azimuth tuning through 360° from 154 AES neurons and showed that azimuth tuning of the AES neurons was usually broad and no systematical change of preferred azimuth was seen.

Neural Codes for Sensory Stimuli

This section reviews two theories on the neural codes for sensory stimuli. One is the traditional view of neural coding and is based on *spike rate*; the other has evolved more recently and incorporates *spike timing* in the theory.

Spike Rate as Neural Codes

Edgar Adrian, who was the first to study the nervous system on the cellular level in 1920s, established three fundamental facts about neural code: (1) individual neurons produce stereotyped action potentials, or spikes; (2) the rate of spiking increases as the stimulus intensity increases; and (3) spike rate begins to decline if a static stimulus is continued for a very long time. Later, the notion of *feature selectivity*, in which the cell's response depends most strongly on a small number of stimulus parameters and is maximal at some optimum value of these parameter, was clearly enunciated by Barlow (1953), who was Adrian's student. A specific example from Barlow's work is the "bug detector" of the frog retina, a class of ganglion cells that respond with great specificity to small black disks moving within neurons' receptive fields (Barlow 1953; also see Lettvin et al. 1959). His "neuron doctrine" formulated from the above observations maintains that sensory neurons are tuned to specific "trigger features" and that a strong discharge by a neuron would signal the presence of a trigger feature within its receptive field (Barlow 1972). In the context of "bug detector," the sensory neurons are represented as yes/no devices, signaling the presence or absence of certain elementary features. As a consequence of this neuron specificity, a given stimulus would be represented by a minimum number of active neurons.

The ideas of feature selectivity and cortical maps have dominated the exploration of the cortex. Cortical map or topographic organization is maintained from sensory epithelia to the sensory cortex. In the visual system, the visual space is mapped to the retina from which a point-to-point projection ascends to the primary visual cortex. The same is true for the somatosensory system in which the sensory input from the body surface projects topographically to the primary somatosensory cortex in the form of a homunculus. In the auditory system, the sensory epithelia in the cochlea is tonotopically organized so that high frequency is represented in the base of the cochlea and low frequency in the apex. Such a tonotopical organization is maintained all the way to the primary auditory cortex.

In other instances, computational maps could emerge from the integrative activity of the central nervous system. For example, many cells in the visual cortex are selective not only for the size of the objects (e.g., the width of a bar) but also for their orientation. Neighboring neurons are tuned to neighboring orientation, so that such a computational feature selectivity is mapped over the surface of the cortex (Hubel and Wiesel 1962). Hubel and Wiesel (1962) also rationalized that this orientation selectivity could be built out of center-surround neurons, suggesting that higher percepts are built out of elementary features. In the auditory system, single neurons in the optic tectum in the barn owl and the superior colliculus in mammals are selective for sound-source location (barn owl: Knudsen 1982; guinea pig: Palmer and King 1982; cat: Middlebrooks and Knudsen 1984; monkey: Jay and Sparks 1984). In those midbrain structures, the preferred sound-source locations of neurons vary systematically according to the

locations of neurons within the structure. In other word, there exists an auditory spatial map in the midbrain.

The neural code based on spike rate leads us quite far in our understanding of the brain function. It is disappointing, however, that despite sustained efforts in several laboratories, a spatial map has not been found in the auditory cortex, a structure essential for sound localization. Previous studies have examined cortical area A1 (Brugge et al. 1994, 1996; Imig et al. 1990; Middlebrooks and Pettigrew 1981; Rajan et al. 1990b), the anterior ectosylvian area (area AES) (Korte and Rauschecker 1993; Middlebrooks et al. 1998) and, to a lesser degree, the anterior auditory field (AAF) (Korte and Rauschecker 1993). Those studies have shown that the spatial tuning of the cortical neurons by spike rate is broad. Moreover, an increased stimulus intensity causes significant expansion of the spatial receptive field in the neurons. At any sound-source location, a stimulus evokes firing from a large proportion of neurons in the auditory cortex (Middlebrooks et al. 1998). There are no systematic shifts in the "best location" of the neurons when the recording electrode changes location in the cortex. The "best location" changes as the stimulus levels are changed. These data are inconsistent with a spike-rate-based topographical code for sound localization. An alternative hypothesis of the neural codes for sound localization, in which spike timing as well as spike counts is incorporated, was proposed and tested by Middlebrooks and colleagues (1994, 1998).

Spike Timing as Neural Codes

As studies of sensory percepts increase in complexity, a simple spike rate code may be rendered inadequate as a predictor of behavior. Although controversy still exists regarding whether spike timing contributes to sensory coding in the cortex (Shadlen and

Newsome 1994; Softky 1995), evidence is rapidly growing that supports the neural codes in which spike timing of the cortical neurons carries information about stimulus parameters. In the context of this review, temporal code is defined as a neural code in which the temporal pattern of a neuron's discharge transmits important information about the stimulus. In the temporal pattern of a neuron's discharge, spike latency and interspike interval enter the picture. Temporal code might also incorporate the relative spike timing among multiple neurons, thus giving rise to the term of ensemble temporal code (Eggermont 1998). Note that a theory of temporal code does not preclude a rate code being superimposed on it simultaneously.

Temporal code has been shown to be superior to rate code in various sensory systems in the following three categories: representation of time-dependent signals, information rates and coding efficiency, and reliability of computation (Rieke et al. 1997). In order for the temporal code to be useful, repetitive firing in the neurons should be sufficiently reliable. Mainen and Sejnowski (1995) demonstrated that the spike-generating mechanisms of the cortical neurons are intrinsically precise. Spike trains could be produced with timing reproducible to less than 1 ms. Such precision is necessary for the propagation of information by a high-resolution temporal code. To address the significance of temporal code, it is necessary to consider not just the intrinsic variability of response to the same stimulus, but also to compare this variability with the variability encountered as stimulus attribute is changed. Victor and Purpura (1996) used a metrical analysis of spike patterns to study the nature and precision of temporal coding in the visual cortex. They found that ~30% of recordings would be regarded as showing a lack of dependence on the stimulus attribute if one considered spike count but

demonstrated substantial tuning when temporal pattern was taken into consideration. Temporal precision was highest for stimulus contrast (10 - 30 ms) and lowest for texture type (100 ms). Their finding suggested the possibility that multiple submodalities can be represented simultaneously in a spike train with some degree of independence. The firing patterns, viewed with high temporal resolution, might represent contrast, while the same pattern, viewed with a substantially lower resolution, might represent texture or another correlate of visual form.

Information about tactile stimulus location is well preserved in the precise topographic maps in the primary somatosensory cortex (SI), as discussed in the previous section. In the secondary somatosensory cortex (SII), neurons have large receptive fields and the topographic organization disappears. Nicolelis and his colleagues (1998) recently showed that different cortical areas could use different combinations of encoding strategies to represent the location of a tactile stimulus. Information about stimulus location could be transformed from a spatial code (based on spike rate) in area SI to an ensemble temporal code in area SII. They made simultaneous multi-site neural ensemble recordings in three areas of the primate somatosensory cortex (areas 3b, SII and 2). An artificial neural network algorithm was then used to measure how well the firing patterns of cortical ensembles could predict, on a single trial basis, the location of a punctate tactile stimulus applied to the animal's body. The neural network could successfully discriminate multiple stimulus locations based on spike patterns of cortical ensembles of each of the three areas. However, by integrating neuronal firing data into a range of bin size (3, 5, 15 or 45 ms), a procedure that was referred to as "bin clumping," they found that the discrimination ability of only area SII neural ensembles was significantly

deteriorated. Therefore, while the neuronal responses in areas 3b and 2 contained information about stimulus location in the form of rate code, the spatiotemporal character of neuronal responses in the SII cortex contained the requisite information using temporally patterned spike sequences (Nicollelis et al. 1998).

Another elegant example of temporal coding comes from reports by Richmond, Optican and their collaborators who used information theory to describe the time dependent neural responses in monkey visual system. The question that they set out to answer was that whether temporal patterns of neuronal firing represent stimulus features such as visual spatial patterns. Their first experiments were done on cells in the inferior temporal cortex (Richmond and Optican 1987), and subsequent experiments have used the same methods to study neurons in several different visual areas (McClurkin et al. 1991; Richmond and Optican 1990). The visual cortical neurons produced the same average number of spikes during the presentation of different spatial patterns (Walsh functions). On the other hand, it was clear that the temporal pattern of spikes during the stimulus presentation was very different (Richmond et al. 1987; 1990). In their studies, they first filtered spike trains in response to a large set of two-dimensional spatial patterns to generate smoothed spike patterns. They then approximated the smoothed spike patterns as a sum of successively more complex waveforms (the principal components). Each instance of the spike pattern was then transformed into a set of coefficients, in much the same way that Fourier series transforms a function of time into the discrete set of Fourier coefficients. It was shown that the first principal component, which was highly correlated with spike count, carried only about half of the information that was available in the spike patterns. Higher principal components, which were

uncorrelated with spike count and yet represented the tendency of the spikes to cluster at different times following the onset of the static visual stimulus, carried nearly half of the total information. Their observations suggested that features of spike patterns additional to spike counts, presumably spike timing, carry stimulus-related information in the visual cortex.

Middlebrooks and collaborators (1994, 1998) showed that spike patterns of auditory cortical neurons carry information about sound-source azimuth. In their studies, an artificial neural network was used as a generic pattern classifier. Such a neural-net algorithm allowed them to "read out" the sound-source azimuth from the firing patterns of single cortical neurons. They observed a moderate level of localization performance based on spike counts alone, and performance improved when spike timing was incorporated. Principal components analysis showed that information-bearing elements of the firing patterns of the cortical neurons included spike counts and temporal dispersion of the firing patterns (Middlebrooks and Xu 1996). Their research along with that of others leads us to the concept of a "panoramic code" in which stimulus-related information is embedded in the temporal patterns of the neuronal discharges. Each single neuron codes many stimulus attributes, e.g., stimulus location around 360° (Middlebrooks et al. 1994; 1998), visual spatial patterns (Richmond et al. 1987; 1990), or visual contrast and texture (Victor and Purpura 1996). With this scheme, one can interpret a continuously varying output of a neuron to decode a continuously varying stimulus parameter. In contrast, a coding scheme based on spike rate would require one to integrate the activity of a neuron over a period of time to obtain a spike rate which is then interpreted as the probability that a particular stimulus is present. In a real-world

situation, the strategy using a timing-based panoramic code is therefore obviously superior to that using a rate-based code in the neural representation of time-dependent sensory information.

CHAPTER 3

SENSITIVITY TO SOUND-SOURCE ELEVATION IN NONTONOTOPIC AUDITORY CORTEX

Introduction

We have shown that the spike patterns of auditory cortical neurons carry information about sound-source azimuth (Middlebrooks et al. 1994, 1998). The principal cues for the location of a sound source in the horizontal dimension (i.e., azimuth) are those provided by the differences in sounds at the two ears, i.e., interaural time difference (ITD) and interaural level difference (ILD). In contrast, the principal cues for location in the vertical dimension are spectral-shape cues that are produced largely by the interaction of the incident sound wave with the convoluted surface of the pinna (see Middlebrooks and Green 1991 for review). The question arises as to whether the spike patterns that we studied represent the output of a system that integrates these multiple cues for sound-source location, or whether they merely demonstrate neuronal sensitivity to an interaural difference that co-varies with sound-source azimuth, such as ILD. Sound sources located anywhere in the vertical midline produce small, perhaps negligible, interaural differences. For that reason, one would predict that a neuron that was sensitive only to interaural differences would show no sensitivity to the vertical location of sound source in the midline and be unable to distinguish front and rear locations. Alternatively, if cortical neurons integrate multiple types of location information, we would expect to observe sensitivity to both the horizontal and the vertical location of a

sound source. We addressed this issue by testing the sensitivity of neurons for the vertical location of sound sources in the median plane.

The spatial tuning properties of cortical auditory neurons have been studied by several groups of investigators (area A1: Brugge et al. 1994, 1996; Imig et al. 1990; Middlebrooks and Pettigrew 1981; Rajan et al. 1990a, 1990b; area AES: Korte and Rauschecker 1993; Middlebrooks et al. 1994, 1998). Most of those studies were restricted to the azimuthal sensitivity of the neurons. Middlebrooks and Pettigrew (1981) described a few units that showed elevation sensitivity to near-threshold sounds, but the stimuli in that study were pure tone bursts, which lacked the spectral information that is crucial for vertical localization of sounds that vary in sound pressure level (SPL). Brugge and colleagues (1994, 1996) confirmed that most A1 cells are differentially sensitive to sound-source direction using "virtual space" clicks as stimuli that simulated 1650 sound-source locations in a three-dimensional space. Near threshold, many of the neurons in their study showed virtual space receptive fields that were restricted in the horizontal and vertical dimensions. When stimulus levels were increased, however, most of the spatial receptive fields enlarged and the vertical selectivity disappeared. Imig et al. (1997) found that, at the level of the medial geniculate body, neurons showed sensitivity to sound-source elevation when stimulated with broadband noise. Such elevation sensitivity disappeared when stimulated with pure tones. They suggested that those neurons were capable of synthesizing their elevation sensitivity by utilizing spectral cues that were present in the broadband noise stimuli.

The present study was undertaken to examine the coding of sound-source elevation by neurons in cortical areas AES and A2. The spike counts of most of these

neurons showed rather broad tuning for sound-source elevation. Nevertheless, spike patterns (i.e., spike counts and spike timing) varied with sound-source elevation. Using an artificial neural network paradigm like the one that we used in the previous studies of azimuth coding (Middlebrooks et al. 1994, 1998), we found that it was possible to identify sound-source elevation by recognizing spike patterns. This result leads us to reject the hypothesis that neurons are merely sensitive to ITD or ILD. Our initial data all were collected from units in area AES (Xu and Middlebrooks 1995). Many of those units failed to discriminate among low elevations. When tested with tones, most of those AES neurons responded only to frequencies greater than 15 kHz. We reasoned that the accuracy in lower elevation coding might improve if we could find neurons that were sensitive to lower frequency tones, because spectral details in the range of 5 to 10 kHz are thought to signal lower elevations (Rice et al. 1992). Therefore, we expanded our experiments to area A2 in which neurons sensitive to broader bands of frequency are more often found. In this report, results from areas AES and A2 were compared in terms of their elevation-coding accuracy and their frequency tuning properties. The role that source sound pressure level might play in elevation coding was addressed. The relationship between network performance in azimuth and elevation of the same neurons was examined.

Methods

Methods of surgical preparation, electrophysiological recording, stimulus presentation, and data analysis were described in detail in Middlebrooks et al. (1998). In brief, 14 cats were used for this study. Cats were anesthetized for surgery with

isoflurane, then were transferred to α -chloralose for single-unit recording. The right auditory cortex was exposed for microelectrode penetration. Our on-line spike discriminator sometimes accepted spikes from more than one unit, so we must note the possibility that we have underestimated the precision of elevation coding by single units. We recorded from the anterior ectosylvian sulcus auditory area (area AES) and auditory area A2. Recordings from area AES were made from the portion of area AES that lies on the posterior bank of the anterior ectosylvian sulcus. Recordings from area A2 were made from the crest of the middle ectosylvian gyrus ventral to area A1. Area A2 was distinguished from neighboring A1 by frequency tuning curves that were at least one octave wide at 40 dB above threshold. Following each experiment, the cat was euthanized and then perfused. The half brain was stored in 10% formalin with 4% sucrose and later transferred to 30% sucrose. Frozen sections stained with cresyl violet were examined with a light microscope to determine the electrode location in the cortex.

Sound stimuli were presented in an anechoic chamber from 14 loudspeakers that were located on the median sagittal plane, from 60° below the frontal horizon (-60°), up and over the head, to 20° below the rear horizon (+200°) in 20° steps. Stimuli consisted of broadband Gaussian noise burst stimuli of 100-ms duration with abrupt onsets and offsets. Loudspeaker frequency responses were closely equalized as described in Middlebrooks et al. (1998). All speakers were 1.2 m from the center of the cat's head. The stimulus levels were 20 to 40 dB above the threshold of each unit in 5-dB steps. A total of 24 to 40 trials was delivered for each combination of stimulus location and stimulus level; locations and levels were varied in a pseudorandom order. Whenever possible, the frequency tuning properties of the units also were studied, using pure tone

stimuli. The pure tone stimuli were 100-ms tone bursts (with 5-ms onset and offset ramps) with frequencies ranging from 3.75 to 30.0 kHz at one-third octave steps. They were presented at 10 dB and 40 dB above threshold from a speaker in the horizontal plane from which strong responses to broadband noise were obtained, usually at contralateral 20 or 40° azimuth.

Off-line, an artificial neural network was used to perform pattern recognition on the neuronal responses (Middlebrooks et al. 1998). Neural spike patterns were represented by estimates of spike density functions based on bootstrap averages of responses to 8 stimuli, as described in the previous paper. The two output units of the neural network produced the sine and cosine of the stimulus elevation, and the arctangent of the two outputs gave a continuously varying output in degree in elevation. We did not constrain the output of the network to any particular range, so the scatter in network estimation of elevation sometimes fell outside the range of locations to which the network was trained (i.e., from -60 to +200°).

Measurement of directional transfer functions of the external ears was carried out in six of the cats after the physiological experiments. A 1/4" tube microphone was inserted in the ear canal through a surgical opening at the posterior base of the pinna. The probe stimuli delivered from each of the 14 speakers in the median plane were pairs of Golay codes (Zhou et al. 1992) that were 81.92 ms in duration. Recordings from the microphone were amplified and then digitized at 100 kHz, yielding a spectral resolution of 12.2 Hz from 0 to 50 kHz. We subtracted from the amplitude spectra a common term that was formed by the root-mean-squared sound pressure averaged across all elevations. Subtraction of the common term left the component of each spectrum that

was specific to each location (Middlebrooks and Green 1990). Those measurements permitted us to study in detail the directional transfer functions of the external ear; however, in the present study, we considered only the spatial patterns of sound levels of three one-octave frequency bands: low-frequency (3.75 - 7.5 kHz), mid-frequency (7.5 - 15 kHz), and high-frequency (15 - 30 kHz).

Results

General Properties of Sound-Source Elevation Sensitivity

A total of 195 units was recorded from areas AES (113 units) and A2 (82 units). Figure 3.1 shows the elevation sensitivity of two AES units (Figure 3.1, A and B) and two A2 units (Figure 3.1, C and D). Left and right columns of the figure plot data from 20 dB and 40 dB above threshold, respectively. The elevation tuning of the units in Figure 3.1, A and C, was among the sharpest in our sample. Most often, however, units showed some selectivity at the lower sound pressure level, but the selectivity broadened considerably at higher sound pressure levels. The units in Figure 3.1, B and D, are typical. The region of stimulus elevation that produced the greatest spike counts from each unit was represented by the "best-elevation centroid", which was the spike-count-weighted center of mass of the peak response, with the peak defined by a spike count greater than 75% of the unit's maximum. The rationale for representing elevation preferences by best-elevation centroids rather than by single peaks or best areas was that the location of a centroid is influenced by all stimuli that produced strong responses, not just by a single stimulus location (Middlebrooks et al. 1998). The primary centroids for the examples in Figure 3.1 are marked by arrows. However, for the responses at 40 dB

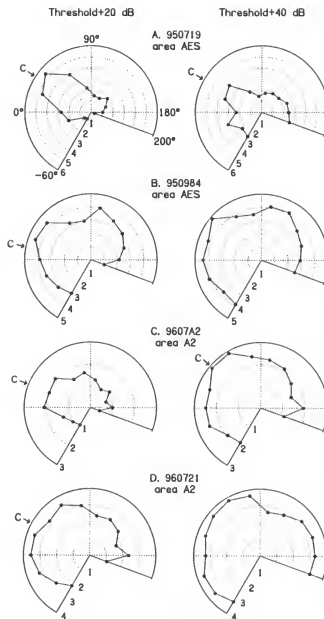


Figure 3.1. Spike-count-versus-elevation profiles. A, B: AES units (950719 and 950984). C, D: A2 units (9607A2 and 960721). The left column represents spike-count-versus elevation profiles at stimulus level 20 dB above threshold and right side 40 dB above threshold. In these polar plots, the angular dimension gives the speaker elevation in the median plane, with 0° straight in front of the cat, 90° straight above the cat's head, and 180° straight behind, as marked in A. The radial dimension gives the mean spike counts (spikes per stimulus presentation). Arrows show the primary elevation centroids, which is the spike-count-weighted center of mass with a peak defined by a spike count greater than 75% of the unit's maximum. No centroids could be calculated for 40 dB data of B and D.

above threshold represented by the right column of Figure 3.1, B and D, no centroids could be computed because the spatial tuning became too flat.

The elevation sensitivity of spike counts in our sample of units is summarized in Figures 3.2 and 3.3. At stimulus levels 20 dB above threshold, 86% of the AES units and 66% of the A2 units showed more than 50% modulation of spike counts by sound-source elevation (Figure 3.2, left panels), but that proportion of the sample dropped to 48% for AES units and 13% for A2 units when the stimulus level was raised to 40 dB above threshold (Figure 3.2, right panels). The height of elevation tuning was represented by the range of elevation over which stimuli activated units to more than 50% of their maximal spike counts. Figure 3.3 shows histograms of the height of elevation tuning, which was defined as the range of elevations over which units responded with spike counts greater than half maximum. Fifty-two percent of the AES units and 84% of the A2 units showed heights larger than 180° at stimulus levels 20 dB above threshold (Figure 3.3, left panels), and the heights of nearly all units from either area AES or area A2 were larger than 180° at 40 dB above threshold (Figure 3.3, right panels). In general, A2 units tended to show broader tuning in sound-source elevation than did AES units (Mann-Whitney U test, $P < 0.01$). Note that all measurements of elevation were made in the vertical midline. Elevation sensitivity might have appeared somewhat sharper if it had been tested in a vertical plane, off the midline that passed through the peaks in units' azimuth profiles. That approach has been used, for instance, in studies of the superior colliculus (Middlebrooks and Knudsen 1984) and medial geniculate body (Imig et al. 1997).

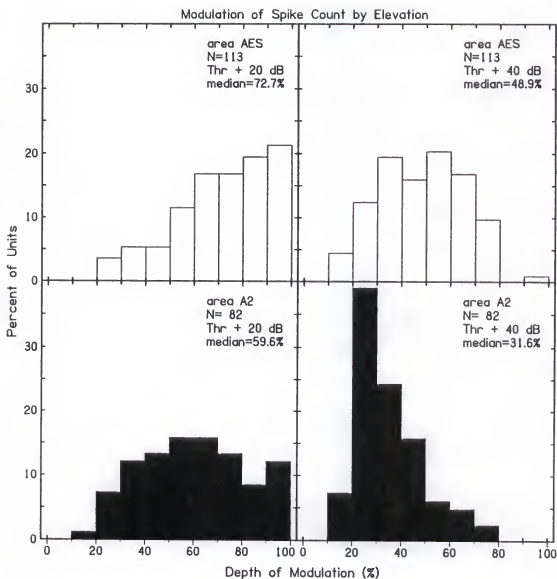


Figure 3.2. Distribution of depth of modulation of spike count by elevation. Open bars in the upper panels represent area AES units. Filled bars in the lower panels represent area A2 units. Left panels plot data at a stimulus level 20 dB above threshold. Right panels plot data at a stimulus levels 40 dB above threshold.

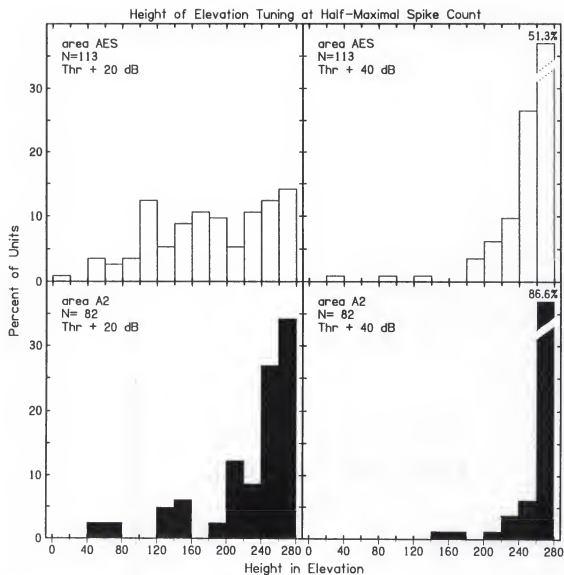


Figure 3.3. Distribution of the range of elevations over which spike counts greater than half maximum were elicited. Conventions as in Figure 3.2.

The best-elevation centroids of our population of 195 units were distributed throughout the elevations of the median plane. However, more centroids were located in the frontal elevations from 20 to 80° than in any other locations (Figure 3.4). For 34% of the AES units and 14% of the A2 units that were studied at 20 dB above threshold, best-elevation centroids were not computed because the modulation of the spike counts of the units by sound-source elevation was smaller than 50%. Such percentages increased to 51 and 87, respectively, at stimulus levels 40 dB above threshold. These units were represented by the bars marked by "NC" in Figure 3.4. No consistent orderly progression of centroids along electrode penetrations was evident in either area AES or area A2. Rarely, for low-intensity stimuli, we saw an orderly progression of centroids along a short distance of the penetration. However, this organization did not persist at higher stimulus levels.

Neural Network Classification of Spike Patterns

Examples of the spike patterns of two AES units and an A2 unit are shown in Figure 3.5 in a raster plot format. Each panel in the figure represents one unit, and only responses elicited at 40 dB above threshold are shown here. Sound-source elevation is plotted on the ordinate and the post-onset time of stimulus is plotted on the abscissa. Each dot represents one spike recorded from the unit. For each of the spike patterns, one can see subtle changes in the numbers and distribution of spikes and in the latencies of the patterns from one elevation to another. It is also noticeable that spike patterns from different units differ significantly.

Figure 3.6 plots the results from artificial neural network analysis of the spike patterns at 40 dB re threshold of the same AES unit as in Figure 3.5A. In panel A,

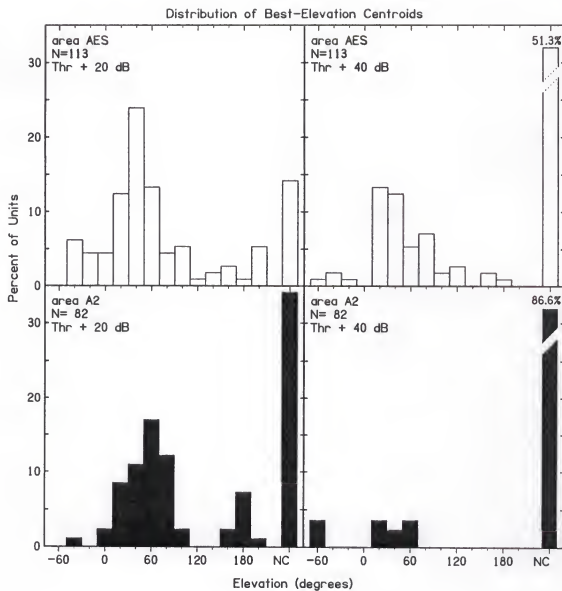


Figure 3.4. Distribution of locations of best-elevation centroids. The percentages of units for which no centroids could be calculated are marked "NC" on the abscissa. Conventions as in Figure 3.2.

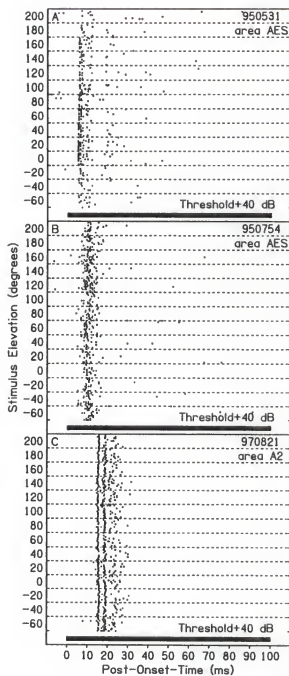


Figure 3.5. Raster plot of responses from two AES units (A: 950531 and B: 950754) and an A2 unit (C: 970821). Each dot represents one spike from the unit. Each row of dots represents the spike pattern recorded from 10 ms before the onset to 10 ms after the offset of one presentation of the stimulus at the location in elevation indicated along the vertical axis. Only 10 of the 40 trials recorded at each elevation are plotted. Stimuli were 100-ms noise burst starting at 0 ms, represented by the thick bars. Stimulus level was 40 dB above threshold.

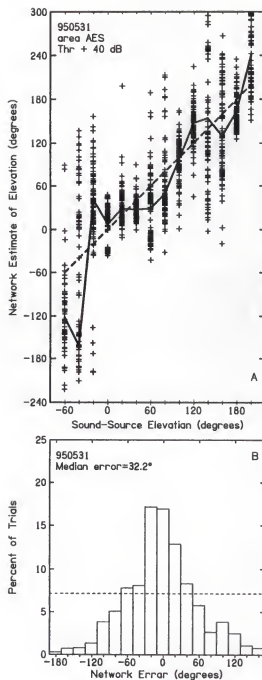


Figure 3.6. Network performance of the same unit (950531) as in Figure 3.5A. In A, each plus sign represents the network output in response to input of one bootstrapped patterns. The abscissa represents the actual stimulus elevation, and the ordinate represents the network estimate of elevation. The solid line connects the mean directions of network estimates for each stimulus location. Perfect performance is represented by the dashed diagonal line. Panel B shows the distribution of network errors. The dashed line represents 7.1%, which is the expected random chance performance given 14 speaker elevations.

each plus sign represents the network estimate of elevation based on one spike pattern, and the solid line indicates the mean direction of responses at each stimulus elevation. In general, the neural-network estimates scattered around the perfect performance line represented by the dashed line. Some large deviations from the targets were seen at certain locations in elevation (e.g., -60 to -20° in this particular example). The neural network classification of the spike patterns of this unit yielded a median error of 32.2° , which was among the smallest in our sample. The distribution of errors in estimation of elevation for this unit is shown in Figure 3.6B. Seventeen percent of network errors were within 10° of the targets. In contrast, the expected value of random chance performance given 14 speakers is 7.1%.

Results of neural-network analysis of responses of another AES unit are shown in Figure 3.7; the spike patterns of this unit are plotted in Figure 3.5B. The network estimates of elevation based on the responses of this unit were less accurate than the estimates shown in Figure 3.6. The network scatter was larger and, at elevations -60 to -20° , the network estimates consistently pointed above the stimuli. Nevertheless, the network produced systematically varying estimates of elevation within the region of 0 to 140° . The unit represented in Figure 3.7 was typical of many units in that network analysis of its spike patterns tended to undershoot elevations at the extremes of the range that we tested (e.g., -60 to -20° and 160 to 200° in this particular example). The median error for this unit was 47.5° , which is slightly larger than the mean of our entire population.

Undershoots at the extremes of the range were also common for A2 units, However, some A2 units could discriminate the lower elevations fairly well. Figure

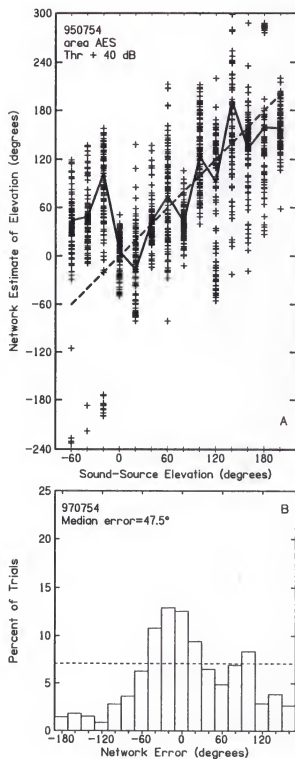


Figure 3.7. Network performance of the same unit (950754) as in Figure 3.5B. Conventions as Figure 3.6.

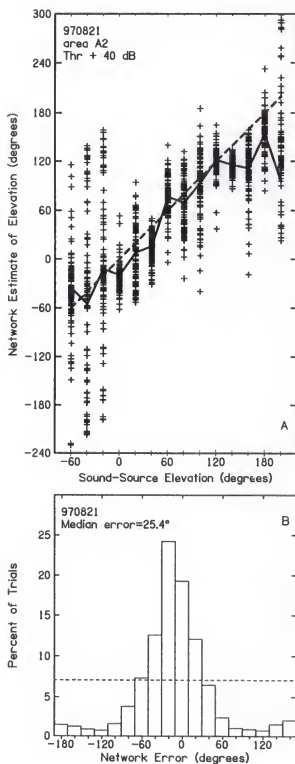


Figure 3.8. Network performance of the same unit (970821) as in Figure 3.5C. Conventions as Figure 3.6.

3.8 shows the network analysis of spike patterns shown in Figure 3.5C. The mean directions of the responses were fairly accurate at all locations except at 160 to 200°, where undershoots were seen (Figure 3.8A). The distribution of errors (Figure 3.8B) shows a bias toward negative errors because of those undershoots.

For all the 195 units studied at 40 dB above threshold, the median errors of the network performance averaged 46.4°, ranging from 25.4 to 67.5°. The distribution of the median errors is shown in Figure 3.9 (right panel). For stimulus level at 20 dB above threshold, the median errors of the network performances averaged 6° less than those at 40 dB above threshold (Figure 3.9, left panel). The bulk of the distribution for all stimulus level conditions was substantially better than chance performance of 65° which is marked by arrows in Figure 3.9. The chance performance of 65° is a theoretical median error when we consider the entire range of 260° of elevation. When we tested the network with data in which the relation between spike patterns and stimulus elevations was randomized, we obtained an averaged median error of $66.5 \pm 1.7^\circ$ across all the 195 units. In general, the median errors of network performance in elevation averaged 2 to 3° larger than those we found in network outputs in azimuth (Middlebrooks et al. 1998). This is consistent with an observation from a study of localization by human listeners (Makous and Middlebrooks 1990). For stimuli in the frontal midline, vertical errors were roughly twice as large as horizontal errors. Results from behavioral studies in cats are difficult to compare in terms of localization accuracy in vertical and horizontal dimensions because only a very limited range of elevation was employed in those studies (Huang and May 1996a; May and Huang 1996).

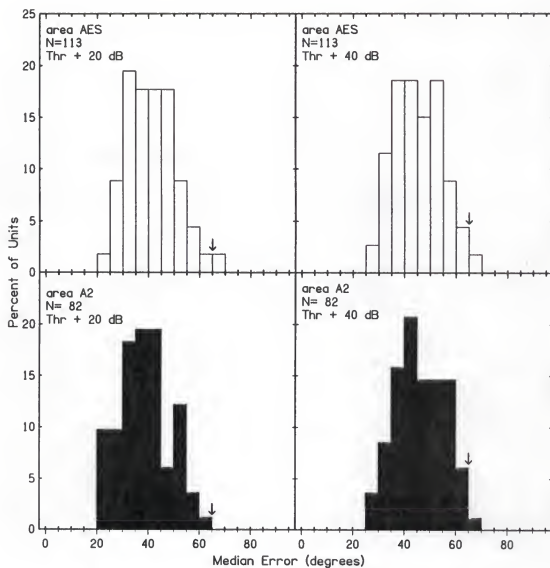


Figure 3.9. Distribution of elevation coding performance across the entire sample of units. Chance performance of 65° is marked by the arrow. Conventions as in Figure 3.2.

We demonstrated in our previous paper that coding of sound-source azimuth by spike patterns is more accurate than coding by spike counts alone (Middlebrooks et al. 1998). We evaluated the coding of sound-source elevation by those two coding schemes. Consistent with our previous paper, we found that median errors in neural network outputs obtained with spike counts were significantly larger than those obtained with complete spike patterns. Median errors in network output obtained in the spike-count-only condition averaged 8 to 12° larger than those obtained in the complete-spike-pattern condition, depending on cortical area (A2 or AES) and stimulus level (20 or 40 dB above threshold).

Comparison of Elevation Coding in Areas AES and A2

We compared our sample of A2 units with our sample of AES units in regard to the accuracy of coding of elevation by spike patterns. Averaged across all elevations, the median errors at sound levels of 20 dB above threshold were slightly smaller for A2 units than those for AES units (t test, $P < 0.05$), but not significantly different from each other in the two areas at 40 dB above threshold (compare upper panels with lower panels in Figure 3.9). When we consider particular ranges of elevation, however, we often found that in area AES, the median errors at locations below the front horizon were much larger than those at the rest of the locations in elevation. In the case of A2 units, this difference was less prominent. Individual examples were given in Figures 3.6 - 3.8. We then calculated the median errors at each of the 14 elevations for units from areas AES and A2. The mean and standard error of the median errors were plotted in Figure 3.10. Asterisks in Figure 3.10 marked the locations at which the differences in the means of the median errors between the two cortical areas were statistically significant (t test, $P <$

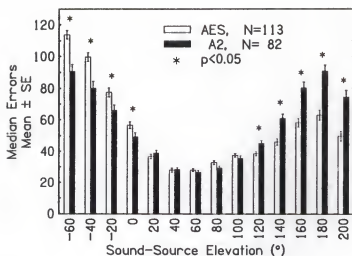


Figure 3.10. Comparison of network performance of A2 and AES units. Plotted here are the means and standard errors of the median errors from the network analysis of AES (open bars) and A2 units (filled bars) at each individual elevation. Asterisks mark the locations where the means of A2 units are significantly different from those of AES units (t test, $P < 0.05$).

0.05). The median errors at elevations from 0 to 120° for A2 units and 20 to 140° for AES units were fairly small. The median errors of AES units at -60 to 0° of elevation were significantly larger than those of A2 units. The reverse was true at 120 to 200° of elevation. Thus, compared to AES units, A2 units achieved a better balance in the network output errors in lower elevations and rear locations.

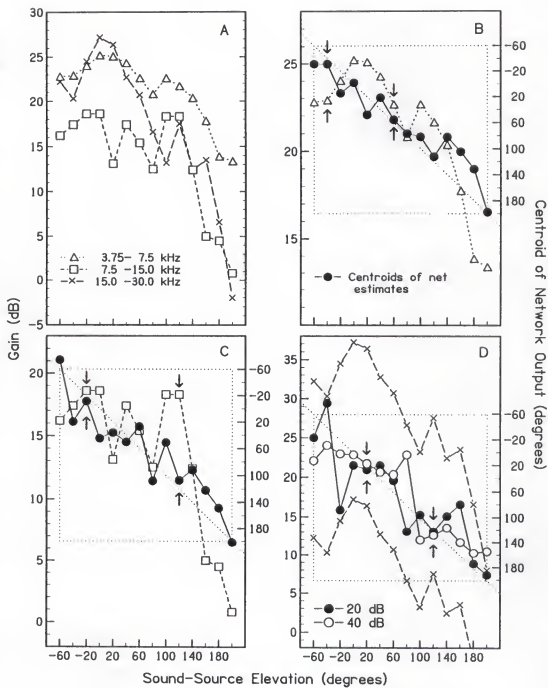
Contribution of SPL Cues to Elevation Coding

Spectral shape cues are regarded as the major acoustical cue for location in the median plane (Middlebrooks and Green 1991). However, the modulation of SPL in the cat's ear canal due to the directionality of the pinna also can serve as a cue. We refer this

cue as the SPL cue. We wished to test the hypothesis that SPL cues alone could account for our results. We measured the SPLs in the cat's ear canal and compared the acoustical data with the network performance. Specifically, we compared the network performance among sound-source elevations at which the stimuli produced similar SPLs in the ear canal. If the SPL cue played a dominant role, the artificial neural network would not be able to discriminate those elevations successfully. We also tested the network performance under conditions in which the SPL of the sound source was varied. If the SPL cue dominated, we would expect that the network performance would be degraded substantially when the variation of the source SPL is large relative to the dynamic range of the modulation of SPL in the cat's ear canal.

The elevation sensitivity of SPLs varies somewhat with frequency, so we measured SPLs within 3 one-octave bands: low, 3.75 - 7.5 kHz; middle, 7.5 - 15 kHz; and high, 15 - 30 kHz. The spatial patterns of sound levels in these three frequency bands were similar among the six cats that were used in the acoustic measurement. Figure 3.11A plots the sound levels in those three frequency bands as a function of sound-source elevation from the measurement of one of the cats. The entire ranges of the sound level profiles for the low-, mid-, and high-frequency regions were 11.9, 17.8, and 29.2 dB, respectively (Figure 3.11A). For the low- and high-frequency bands, sound from 0° elevation produced the maximal gain in the external ear canal of the cat. Sound levels decreased more or less monotonically when the sound source moved below or above the horizontal plane and behind the cat. For the mid-frequency band, however, sounds from -20 and 0° and those from 100 and 120° produced the largest gains in the

Figure 3.11. Sound levels and neural network performance. A: Sound levels measured at the external ear canal as a function of sound-source elevation. Levels were measured in low- (3.75 - 7.5 kHz), mid- (7.5 - 15 kHz), and high-frequency (15 - 30 kHz) bands. B: Sound levels in the low-frequency band are plotted with triangles on the left ordinate. The mean directions of neural network responses of a unit (960553) that responded well to the low-frequency tones are plotted with filled circles on the right ordinate. The two ordinates are scaled so that the ranges of two curves roughly overlap. The small arrows mark the pair of sound-source elevations at which sound levels were found similar to one another (within 1 dB) but at which network estimates of elevation were different. C: Sound-level profile at mid-frequency region (open squares) and mean directions of the network responses (filled circles) of a unit (950915) that responded well to mid-frequency tones are plotted in the same format as B. D: Sound-level profiles at high-frequency band at 10 dB above and 10 dB below the actual one shown in A are plotted on the left ordinate with crosses to simulate the 20-dB range of the roving levels. Mean directions of the network responses of a unit (950702) that responded well to high-frequency tones are plotted on the right ordinate. The network was trained with spike patterns from 5 SPLs, from 20 to 40 dB above threshold. Filled and open circles are mean directions of network output when tested with spike patterns obtained with stimulus at 20 and 40 dB above threshold. Arrows mark examples at which the two network outputs point to the same correct locations.



external ear canal. The sound levels dropped at locations behind the cat and in those below the frontal horizon.

We compared the elevation sensitivity of sound levels with the neural network estimation of elevation by plotting sound levels and neural network output on common abscissas (Figure 3.11, B and C). Figure 3.11B shows the network analysis of a unit that responded best to frequencies in the low-frequency band. The triangles show the sound levels in that band. Figure 3.11C shows network data and mid-frequency sound levels for a unit that responded best to the middle frequencies. The left ordinate, used for SPL data, and the right ordinate, used for neural network estimate, were scaled so that both sets of data roughly overlapped. If the network identification of elevation was due simply to SPL variation, sound sources that differed in elevation but produced the same SPLs in the ear canal would result in the same elevations in the network output. In fact, the neural network could distinguish pairs of speakers at which similar SPLs (within 1-dB) were produced. Examples of such pairs of locations are marked by arrows in Figure 3.11, B and C. The results are inconsistent with the prediction based on the SPL cue.

Next, we tested the effect of roving the source SPLs. Figure 3.11D was plotted for another unit in a similar format to Figure 3.11, B and C. This unit responded best to frequencies in the high-frequency band. Here, we plotted two high-frequency sound-level curves separated by 20 dB, simulating the SPL cues under conditions in which we varied the stimulus SPLs in a range of 20 dB. A neural network was trained with spike patterns from five SPLs between 20 and 40 dB above threshold in 5-dB steps. The network output based on spike patterns elicited with single source SPLs at 20 and 40 dB above threshold were plotted using the right ordinate. One can see from Figure 3.11D

that even though the high-frequency band provided the strongest SPL cues for localization in elevation, those SPL cues were greatly confounded when stimulus levels were roved in the range of 20 dB. For instance, a stimulus of 20 dB SPL at 0° and a stimulus of 40 dB SPL at 180° would produce similar sound level at the ear canal. Nevertheless, neural-network recognition of spike patterns produced by two single stimulus levels (20 and 40 dB above threshold) were fairly accurate and comparable. Arrows show examples in which the network recognized two sets of spike patterns as responses to stimuli at the same elevation, even when the stimulus SPLs differed by 20 dB. The median error in network output for the unit represented in Figure 3.11D was 29.0°. That means that one half of the network outputs fell within a range of roughly 58.0° ($\pm 29.0^\circ$) around the correct elevation. That range of errors is 22.3% of the 260° range of elevation that was tested. In contrast, SPL cues to sound-source elevation were confounded by source levels that roved over a range of 20 dB, which is 68.5% of the 29.2-dB range of variation of SPL produced by a constant-level source moved through 260° of elevation. We applied the same approach as in Figure 3.11 to all the units in our sample that had median errors smaller than 40° and obtained results qualitatively similar to those shown in the figure. These results contradict the hypothesis that elevation sensitivity is due entirely to the elevation dependence of SPL.

Our systematic analysis of the effect of roving levels on network performance further supports the hypothesis that level-invariant information about sound-source location is present in the spike patterns. For the sample of 195 units, the averaged median errors of the network when trained and tested with responses to stimuli that were 20 and 40 dB above threshold were 40.3 and 46.4°, respectively. Neural network

analysis yielded an average median error of 47.9° when trained and tested with 5 roving levels (20, 25, 30, 35, and 40 dB above threshold). Statistics did not show any significant difference of the averaged median errors between the condition of a single level at 40 dB above threshold and that of 5 roving levels (paired t test, $P > 0.05$).

Frequency Tuning Properties and Network Performance

The coding of sound source elevation requires integration of information across a range of frequencies. Frequency tuning properties of a neuron might be related to a neuron's elevation sensitivity. In this section, we explored the relation between the frequency tuning properties and the network performance in the two cortical areas. We found that A2 units showed broader frequency tuning than did AES units. The broader frequency tuning in A2 was mainly due to that the low-cutoff frequencies of the frequency tuning curves of the A2 units extended toward lower frequencies. Acoustic measures of the cat's head-related transfer function (Rice et al. 1992) and behavioral studies in cats (Huang and May 1996a) suggested that spectral details in lower frequency range (e.g., 5 - 10 kHz) might signal low elevations. In fact, as we showed earlier, the AES units tended to produce larger errors in the low elevations (-60 to 0°) than did A2 units (Figure 3.10). Could the broader frequency tuning and lower low-cutoff frequencies of the A2 units account for their better performance in the low elevations?

First, we consider the frequency tuning properties of the units. The units that we encountered in areas AES and A2 responded well to broadband noise burst stimuli. We recorded frequency tuning responses to tone bursts of 100-ms duration in 173 of the 195 units. Among them, 91 units were from area AES and 82 from area A2. Most of units showed stronger responses to higher frequency tones (≥ 15 kHz) than to lower frequency

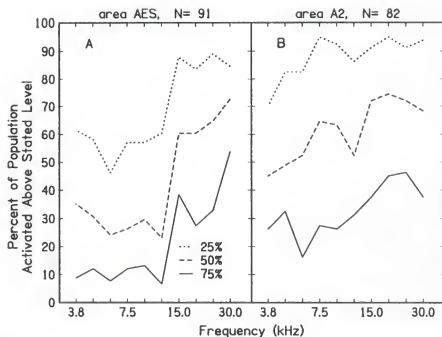


Figure 3.12. Percentage of unit sample activated as a function of stimulus tonal frequency. The three lines in each panel represent the percentage of units activated at or above 25, 50, and 75% of maximal spike counts. A. Pooled data from 91 AES units. B. Pooled data from 82 A2 units.

tones (<15 kHz). Figure 3.12, A and B, shows, for our sample of AES and A2 units, respectively, the percentage of the population activated to levels at or above 25, 50, and 75% of maximal spike counts at various tonal frequencies, at a stimulus level 40 dB above threshold. At almost all frequencies, more than half of the population in both areas AES and A2 were activated above 25% of maximal spike counts. Tonal stimuli activated a larger fraction of the unit population in area A2 than in area AES, especially in lower frequencies. Hence, frequency tuning bandwidth appeared broader in our sample of A2

units than in the AES units. The conventional way of defining tuning bandwidth is to find thresholds at various frequencies and then to measure the bandwidth at a certain level above the lowest threshold. That might not provide an accurate description of tuning bandwidth under condition of free-field sound stimulation because the transfer functions of the pinnae will be added to the frequency sensitivity of the unit. Instead, we defined the tuning bandwidth as follows. First, we measured spike counts in response to tones at various frequencies with a fixed level of 40 dB above the threshold for the best frequency. The tuning bandwidth was the frequency range over which the spike counts were at or above 50% of the maximal spike count. That provided a somewhat more appropriate measure of the bandwidth of frequency that influenced the unit responses in our study. The distribution of the frequency tuning bandwidths in our sample of A2 and AES units is shown in the upper panels of Figure 3.13. The mean bandwidth in A2 was 2.02 octaves and that in AES neurons was 1.49 octaves. This difference was statistically significant (t test, $P < 0.01$).

Next, in order to explore whether this difference in frequency tuning bandwidth could account for the difference between AES and A2 units in neural network performance in low elevation coding, we measured the correlation of the bandwidths of individual A2 and AES units with their neural network performance, particularly in the lower elevation coding. Lower panels of Figure 3.13 are scatter plots of the neural network performance at lower elevations as a function of frequency tuning bandwidth for our AES and A2 units, respectively. The lower elevations that represented are -60 to 0° , which are in the range in which difference between the two cortical areas were evident (Figure 3.10). No correlation could be seen between the network performance

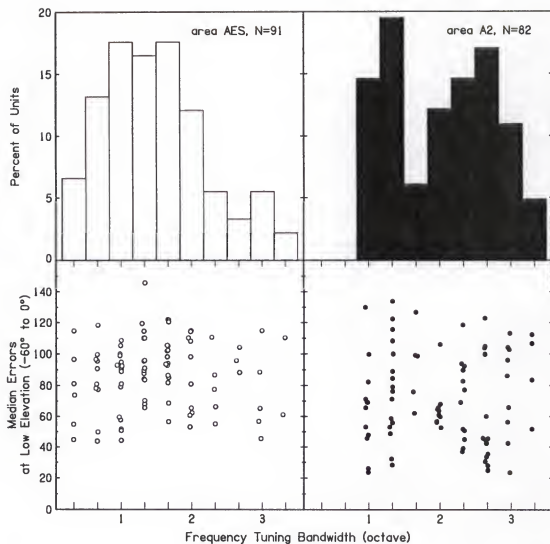


Figure 3.13. Frequency tuning bandwidth and neural network performance. Upper panels represent the distribution of bandwidth in AES units (left, open bars) and in A2 units (right, filled bar). Lower panels represent relation between the neural network performance in the lower elevation and the frequency tuning bandwidth. Left and right panels represent areas AES and A2, respectively. Median errors were computed in a range of -60 to 0° elevation.

represented by the median errors and the frequency tuning bandwidth. Similarly, we measured the correlation of the low-cutoff frequencies of the frequency tuning curves of individual A2 and AES units with their neural network performance in the lower elevations. We found a marginally significant correlation between the network output errors at low elevations and low-cutoff frequencies in the sample of A2 units ($r = 0.24$, $0.01 < P < 0.05$) but not in the sample of AES units.

Relation between Azimuth and Elevation Coding

For 175 units, responses to stimuli from both horizontal and vertical speakers were obtained. Across these 175 units, there was a significant positive correlation between the network performance in azimuth and in elevation (Figure 3.14). Each panel in Figure 3.14 is a scatter plot of the median errors of the same units in encoding sound-source azimuth and elevation. AES units ($N=113$) are presented in the upper panels and A2 units ($N=62$) in the lower panels. Left panels plot data obtained from stimulus level at 20 dB above threshold and right panels 40 dB above threshold. Correlation coefficients (r) between median errors in azimuth and elevation ranged between 0.23 to 0.53 depending on the cortical areas and the stimulus levels. The correlation coefficients of the A2 units were larger than those of the AES units, especially for the stimulus level at 40 dB above threshold. Among the units that coded elevation with median errors of 40° or less, for example, the majority of units also showed median errors of 40° or less in azimuth. The principal acoustic cues for localization in elevation differ from those for localization in azimuth. If neurons are sensitive only to a particular localization cue, no correlation or perhaps negative correlation between network performance in the two dimensions would be expected. The fact that we observed positive correlations between

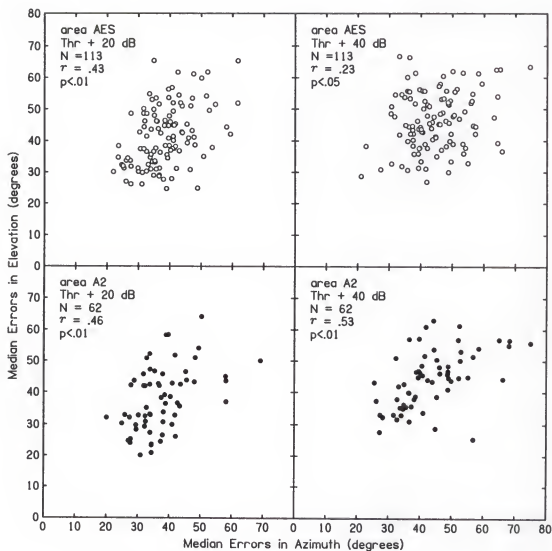


Figure 3.14. Correlation between network performance in azimuth and elevation. Each dot in the scatter plots represents, for one unit, the median error of the network performance in elevation versus that in azimuth. There is a positive correlation between network performance in both dimensions. Open circles in the upper panels represent area AES units. Filled circles in the lower panels represent area A2 units. Left panels plot data at a stimulus level 20 dB above threshold. Right panels plot data at a stimulus level 40 dB above threshold.

the two dimensions indicates that many units can integrate information from multiple types of localization cues.

Discussion

Results presented in Middlebrooks et al. (1998) support the hypothesis that sound-source azimuth is represented in the auditory cortex by a distributed code. In that code, responses of individual neurons carry information about 360° of azimuth, and the information about any particular sound-source location is distributed among units throughout entire cortical areas. The present study extends that observation to the dimension of sound-source elevation. The acoustical cues for sound-source elevation differ from those for azimuth, and identification of source azimuth and elevation presumably require distinct neural mechanisms. The observation that units in areas AES and A2 show similar coding for azimuth and elevation supports the hypothesis that neurons integrate the multiple cues that signal the location of a sound source rather than merely coding a particular acoustical parameter that happens to co-vary with sound-source location. In this Discussion, we consider the acoustical cues that could underlie the elevation sensitivity that we observed, evaluate the similarities and differences between areas AES and A2 in regard to elevation and frequency sensitivity, and comment on the significance of the correlation between azimuth and elevation coding accuracy.

Acoustical Cues and Localization in Median Plane

Acoustical measurements of directional transfer functions in the ear canal and behavioral studies have provided insights into the acoustical cues for sound localization in the vertical dimension. Due to the approximate left-right symmetry of the head and

ears, a stimulus presented in the median plane will reach both ears simultaneously with equal levels. Interaural time differences and interaural level differences that are important for localization in the horizontal plane may contribute little if any to the localization in the median plane (Middlebrooks and Green 1991; Middlebrooks et al. 1989).

Sound pressure level, on the other hand, can be a cue for vertical localization if the source level is known and constant. The SPL in the ear canal varies with sound-source elevation. Earlier recordings in cats have shown that within the range of -60 to $+90^\circ$ elevation, SPL varies a few dB for lower frequency tones to as much as 20 dB for high frequency tones (Middlebrooks and Pettigrew 1981; Musicant et al. 1990; Phillips et al. 1982). In the present study, the acoustical recording of the directional transfer function at the entrance of the external ear canal of cats was carried out in the range of elevation from -60 to 200° . Instead of examining each individual frequency, we plotted the SPL profile in three frequency bands (Figure 3.11A). The high-frequency band (15 - 30 kHz) had the largest variation in SPL. The entire range of the sound level profiles for the low-, mid-, and high-frequency regions were 11.9, 17.8, and 29.2 dB, respectively. To test the degree to which SPL cues might have contributed to our physiological results, we compared the elevation sensitivity of unit responses with the elevation sensitivity of ear-canal SPLs. There were two indications that SPL cues are not the principal cues for the elevation sensitivity we observed. First, we observed many instances in which sound sources at two locations produced roughly the same SPL in the ear canals, yet produced unit responses that could be readily distinguished by an artificial neural network. Second, under conditions in which we roved stimulus SPLs over a range of 20 dB, a sound source at a single location produced SPLs ranging over 20 dB, yet

produced unit responses containing SPL-invariant features that resulted in roughly equal neural-network estimates of elevation. Although SPL cues might contribute to elevation sensitivity under certain conditions in which sound-source SPLs are constant, these two observations indicate that SPL cues alone could not have accounted for the neuronal elevation sensitivity that we observed.

A body of evidence suggests that spectral-shape cues are the principal cues for localization in the vertical dimension. Measurement of the directional transfer functions of human ears (Middlebrooks et al. 1989; Shaw 1974; Wightman and Kistler 1989) and those of cat ears (Musicant et al. 1990; Rice et al. 1992) has shown that spectral shape features vary systematically with sound-source elevations. The most conspicuous features of the transfer functions of a cat ear are probably the spectral notches. The center frequencies of the spectral notches (5-18 kHz in cat) increase as sound-source elevation changes from low to high (Musicant et al. 1990; Rice et al. 1992). Recent behavioral studies in cats have provided evidence that indicates that the mid-frequency spectral-shape cues are important for vertical localization (Huang and May 1996a, 1996b; May and Huang 1996). A recent report from Imig and colleagues (1997) has demonstrated that at least some elevation sensitive units in the medial geniculate body lose that sensitivity when tested with tonal stimuli, also suggesting a spectral basis for elevation sensitivity (Imig et al. 1997). We do not yet have any direct evidence that the elevation sensitivity that we observed was due to sensitivity to spectral-shape cues. Having ruled out SPL cues, however, sensitivity to spectral-shape cues certainly is the most likely explanation for the elevation sensitivity that we see.

A2 versus AES: Elevation Sensitivity and Frequency Tuning Properties

Our initial data from area AES showed larger errors at frontal locations below the horizon than at higher elevations and in the rear. We explored auditory area A2 to test whether sensitivity to low frontal elevations might be more accurate in another cortical area. Averaged across all elevations, the accuracy of elevation coding for units from areas A2 and AES was not significantly different. Nevertheless, differences between cortical areas were found in the errors at low frontal and rear locations (i.e., -60 to 0° and $+120$ to $+200^\circ$). For both cortical areas, errors of the network output at lower elevations and rear locations were much larger than those at other locations. These large errors were almost always caused by underestimation of targets. These undershoots might be due to an edge effect of the neural network analysis. That is, the network would tend not to give mean outputs at locations beyond the limits of the training set. However, the edge effect could not explain why there were differences in the accuracy of network output in various elevation ranges between the two cortical areas.

Since spectral-shape cues of the sound are important for localization in vertical plane, it is conceivable that differences in the frequency tuning of neurons in areas AES and A2 might account for differences in elevation sensitivity. Previous studies showed that broadly tuned neurons were found in both areas (Andersen et al. 1980; Clarey and Irvine 1986; Reale and Imig 1980; Schreiner and Cynader 1984). In area AES, neurons were shown to respond to ranges of frequency that most often were weighted toward high frequencies (Clarey and Irvine 1986). In area A2, a dorsoventral gradient of frequency tuning bandwidth was demonstrated with the lowest Q_{10} values found in the most ventral parts of A2. Frequency bands often extended to low frequencies (Schreiner

and Cynader 1984). For the sample of our 91 AES units and 82 A2 units, most of them showed stronger responses to higher frequency tones (≥ 15 kHz) than to lower frequency tones (< 15 kHz). Frequency tuning bandwidth was broader in our sample of A2 units than in the AES units, and tonal stimuli activated a larger fraction of the unit population in area A2 than in area AES, especially at lower frequencies (Figures 3.12 and 3.13). We could postulate that the properties of broad frequency tuning in area A2 would make A2 neurons more suitable for detecting the spectral shape cues that are important for elevation coding than AES neurons. However, our results were not conclusive in this regard. No correlation was found between the frequency tuning bandwidth and the network output errors at the locations at which differences between A2 and AES neurons were evident (Figure 3.13). Only a marginally significant correlation was found between the low-cutoff frequencies and network output errors at low elevations in the sample of A2 units. Perhaps overall frequency tuning bandwidth of the cortical neurons is not as important as are details of frequency response areas that consist of excitatory and inhibitory regions, as suggested in the data obtained from the medial geniculate body (Imig et al. 1997). Our limited data, as well as earlier studies on frequency tuning of the A2 and AES neurons, have shown that some of the neurons from either cortical area have irregular frequency tuning curves in which two or multiple peaks are present (Clarey and Irvine 1986; Schreiner and Cynader 1984). Such irregular frequency tuning may produce spectral regions of inhibition and facilitation which in turn may provide the basis for a neuron's directional sensitivity.

Correlation between Azimuth and Elevation Coding

We find that, in general, those cortical units in areas AES and A2 that exhibit the most accurate elevation coding also tend to show good azimuth sensitivity. The psychophysical literature supports the view that azimuth sensitivity derives primarily from interaural difference cues and that elevation sensitivity derives from spectral shape cues (Middlebrooks and Green 1991). We would like to conclude that single cortical neurons receive information both from brain systems that perform interaural comparisons as well as those that analyze details of spectra at each ear. An alternative interpretation, however, is that the units that we studied were not sensitive to interaural differences and that both the azimuth sensitivity and the elevation sensitivity that we observed were derived from spectral shape cues. Indeed, acoustical studies in cat and human indicate that spectra measured at each ear vary conspicuously as a broadband sound source is varied in azimuth (Rice et al. 1992; Shaw 1974). Moreover, human patients that are chronically deaf in one ear can show reasonably accurate localization in azimuth, presumably by exploiting monaural spectral cues for azimuth (Slattery and Middlebrooks 1994).

These conflicting conclusions can be resolved only by future studies in which specific acoustical cues are controlled directly. At this time, however, at least two lines of evidence lead us to reject the view that the spatial sensitivity of the units that we studied is derived entirely from spectral shape cues. *First*, Imig and colleagues (1997) searched for units in the cat's medial geniculate body that showed azimuth sensitivity derived predominantly from monaural spectral cues. Only about 17% of units in the ventral nucleus (VN) and the lateral part of the posterior group (PO) showed azimuth

sensitivity that persisted after the ipsilateral ear was plugged. That study is not directly relevant to the current one, since VN and PO project most strongly to cortical area A1, not A2 or AES. Nevertheless, those results argue that in at least two divisions of the auditory thalamus only a small minority of units shows azimuth sensitivity that is dominated by monaural spectral cues. *Second*, studies in area A2 that used dichotic stimulation have shown that about a third of area A2 units show excitatory/inhibitory binaural interactions (Schreiner and Cynader 1984). That type of binaural interaction would necessarily result in sensitivity to interaural level differences. About 40% of units in area A2 and ~69% of units in area AES show excitatory/excitatory binaural interactions (Clarey and Irvine 1986; Schreiner and Cynader 1984), and excitatory/excitatory interactions also can result in sensitivity to interaural level differences (Wise and Irvine 1984). Even if we consider only the excitatory/inhibitory units in area A2, a minimum of a third of our A2 sample should have included units that were sensitive to interaural level differences. It would be difficult to argue that both the elevation and azimuth sensitivity shown by units in areas AES and A2 is due primarily to spectral shape sensitivity.

Concluding Remarks

The study reported in Middlebrooks et al. (1998) demonstrated that the responses of single units in areas AES and A2 can code sound-source location in the horizontal plane throughout 360° of azimuth. That result raised the question of whether units in those cortical areas integrate multiple acoustical cues for sound-source location or whether they simply code the value of a single acoustical parameter, such as interaural level difference, that co-varies with azimuth. In the present study, we have found that

the responses of units also can code the elevation of a sound source in the median plane, in which interaural difference cues presumably are negligible. Moreover, the units that show the best elevation coding accuracy also code azimuth well. These results do not constitute conclusive evidence of a direct role of these neurons in sound-localization behavior. They do, however, support the hypothesis that single cortical neurons can combine information from multiple acoustical cues to identify the location of a sound source in azimuth and elevation.

CHAPTER 4

AUDITORY CORTICAL SENSITIVITY TO VERTICAL SOURCE LOCATION: PARALLELS TO HUMAN PSYCHOPHYSICS

Introduction

We have reported previously that the spike patterns (spike counts and spike timing) of neurons in the nontonotopic auditory cortex carry information about sound-source location (Middlebrooks et al. 1994, 1998; Xu et al. 1998). The results support the hypothesis that the activity of individual neurons carries information about broad ranges of location and that accurate sound localization is derived from information that is distributed across large population of neurons. The spike patterns that we studied represent an output of a system that integrates multiple cues for sound-source location.

Human psychophysical studies have demonstrated that accurate localization of broadband sounds in the vertical plane utilizes spectral-shape cues that are produced by the interaction of the incident sound wave with the head and the convoluted surface of the pinna (see Middlebrooks and Green 1991 for review). Human listeners can localize accurately when presented with stimuli that have spectra that are fairly broad and flat, as is true of most natural sounds. When certain filters are applied to stimuli, however, localization based on spectral shape cues is confounded and listeners make systematic errors in the vertical and front/back dimensions. Similarly, behavioral studies in cats have shown that cats can accurately localize broadband sounds in the vertical plane and that

vertical localization fails when stimulus spectra are restricted to narrow bands of frequency (Huang and May 1996a; May and Huang 1996; Populin and Yin 1998).

If the neurons that we have studied in the auditory cortex contribute to sound localization behavior, one would expect that their responses would correctly signal the locations of broadband sound sources, as we have observed previously. By analogy with behavioral results, we also would expect their responses to signal systematically incorrect locations when presented with certain filtered sounds. It is that expectation that we tested in the present study.

We chose to study auditory cortical area A2 because A2 neurons are broadly tuned to frequency (Andersen et al. 1980; Reale and Imig 1980; Schreiner and Cynader 1984) and because elevation sensitivity encoded by their spike patterns has been shown in the previous report (Xu et al. 1998). Stimuli consisted of broadband noise and three types of filtered noise. Broadband noise was chosen because human and feline listeners tend to localize sounds accurately in the vertical and front/back dimensions when stimulus spectra are broad and flat (Makous and Middlebrooks 1990; May and Huang 1996). The filtered noise included narrow bandpass noise (narrowband noise), narrow band-reject noise (notched noise) and highpass noise. We chose narrowband noise because human listeners make systematic errors when required to localize a narrowband sound and because that pattern of errors is predicted well by a quantitative model (Middlebrooks 1992). Similar behavioral results were observed in a head-orientation experiments in cats (Huang and May 1996a). We chose notch stimuli because a possible localization illusion due to spectral notches was observed in a human behavioral studies (Bloom 1977; Walkins 1978) and because analysis of feline head-related transfer

functions has led several groups to speculate that notches might provide salient cues for localization (Musicant et al. 1990; Rice et al. 1992). Highpass noise was chosen because behavioral studies have shown that human localization judgements are influenced by the cut-off frequency of a highpass sound (Hebrank and Wright 1974b) and because recent human psychophysical studies from this laboratory have shown that narrowband and highpass noise stimuli that have equal low-frequency cut-offs tend to produce equivalent localization judgements (Macpherson and Middlebrooks 1999).

In the present study, we performed pattern recognition on cortical spike patterns using an artificial neural network paradigm that we employed in previous studies of azimuth and elevation coding (Middlebrooks et al. 1994, 1998; Xu et al. 1998). We trained neural networks to recognize the spike patterns elicited by broadband noise sources at various elevations. When presented with such spike patterns, the trained networks produced estimates of the source location that corresponded reasonably well with the actual locations. Later, the trained network was used to classify cortical responses to filtered noise. In response to spike patterns elicited by narrowband noise of a given center frequency, the network produced fairly constant elevation estimates, regardless of the actual source elevation. When presented with spike patterns elicited by narrowband sounds that varied in center frequency, the network produced elevation estimates that tended to vary systematically in elevation. The region in elevation that was associated with a given center frequency could be predicted by a localization model based on spectral shape recognition. Highpass stimuli tend to produce spike patterns and network outputs similar to those of narrowband stimuli when the low-frequency cut-offs of both stimuli match each other. Our data support the hypothesis that the elevation

sensitivity of these cortical neurons derives from computational principles similar to those that underlie human vertical localization.

Methods

Eight adult cats of either sex were used in this study. Cats were anesthetized for surgery with isoflurane, then were transferred to α -chloralose for single-unit recording. The right auditory cortex was exposed for microelectrode penetration. Both ears of the cat were supported in a symmetrical forward position that resembled the ear position adopted by a cat attending to a frontal sound. Details of anesthesia procedures and surgical preparation are available in Middlebrooks et al. (1998).

Experimental Apparatus

Experiments were conducted in a sound-attenuating chamber that was lined with acoustical foam (Illbruck) to suppress reflections of sounds at frequencies > 500 Hz. Sound stimuli were presented from loudspeakers (Pioneer model TS-879 two-way coaxials) mounted on 2 circular hoops, one in the horizontal plane and one in the vertical midline plane. On the horizontal hoop, 18 loudspeakers spaced by 20° covered 360° . On the vertical hoop, 14 loudspeakers spaced by 20° ranged from 60° below the frontal horizon, up and over the top, to 20° below the rear horizon. Vertical locations were labeled continuously in 20° steps from -60 to 200° . All loudspeakers had a distance of 1.2 m from the center of the chamber where the head of the animal was positioned. In the present study, we focused only on the vertical plane.

Experiments were controlled with an Intel-based personal computer. Acoustic stimuli were synthesized digitally, using equipment from Tucker-Davis Technologies

(TDT). The sampling rate for audio output was 100 kHz, with 16-bit resolution. Before each experiment, the loudspeakers were calibrated by presenting maximum-length sequences (Golay codes) and recording the responses with a 1/2-in microphone (Larson-Davis model 2540) placed in the center of the chamber in the absence of the cat (Golay 1961; Zhou et al. 1992). Loudspeaker responses were equalized individually so that the root-mean-squared variation in sound level, computed in 6.1-Hz steps from 1,000 to 30,000 Hz, was < 1.0 dB.

Multichannel Recording and Spike Sorting

We used silicon-substrate thin-film multichannel recording probes to record unit activities. Each probe had 16 recording sites on a one-dimensional shank spaced at intervals of 100 μm and allowed simultaneously recording from up to 16 sites (Drake et al. 1988; Najafi et al. 1985). The nominal impedances were $\sim 4\text{ M}\Omega$. We recorded from auditory cortical area A2. The probe was passed in a dorsoventral orientation, roughly parallel to the cortical surface, near the crest of the ventral middle ectosylvian gyrus. Generally, the probe passed through the middle cortical layers that are active under anesthesia, although recordings did not necessarily all come from the same cortical layer. An on-line spike discriminator (TDT model SD1) and custom graphic software were used to monitor spike activities from one selected channel at a time. Prior to detailed study at each probe placement, we determined the frequency tuning properties of units at the most dorsal recording sites. We sometimes detected sharp frequency tuning, which was taken as evidence that the probe was in the auditory cortical area A1. In such cases, we retracted the probe and moved it further ventral.

Signals from the recording probe were amplified with a custom 16-channel amplifier, digitized at a 25-kHz rate, sharply low-pass filtered below 6 kHz, re-sampled at a 12.5 kHz sample rate, and then stored on a PC hard disk. Off-line, we isolated unit activities from the digitized signal using custom spike-sorting software. Spike times were stored at 20- μ s resolution for further analysis. Occasionally, we encountered well-isolated single units, but most often the recordings were characteristic of unresolved clusters of several units. We presume that the addition of responses of multiple units could only increase the apparent breadth of spatial tuning of single units and could only decrease the spatial specificity of spike patterns. For that reason, we regard our results to be conservative estimates of the accuracy of spatial coding by single units. Some unit recordings were regarded as weak or unstable and thus were excluded from further analysis. Usable recordings met the following two criteria. (1) In response to broadband noise, the maximum mean spike rate across all tested sound levels and elevations was ≥ 1 spike per trial. (2) Across all presentations of broadband noise, the mean spike rate in the first half of the trials differed from that in the second half by no more than a factor of 2.

Stimulus Paradigm and Experimental Procedure

At each placement of a recording probe, we recorded responses to tones, broadband noise, and filtered noise. The entire stimulus set required about 6-8 hours to present. We first studied the frequency tuning properties of the units. Pure tone stimuli, consisted of 80-ms tone bursts (with 5-ms onset and offset ramps) with frequencies ranging from 1.18 to 30.0 kHz in 1/3-oct steps. They were presented at +80 or +100°

elevation at stimulus levels of 10, 20, 30 and 40 dB above the threshold of the most sensitive unit.

Elevation sensitivity was then studied by presenting broadband noise bursts from the 14 loudspeakers in the vertical midline plane, one loudspeaker at a time. The broadband noise stimuli consisted of independent Gaussian noise samples of 80-ms duration (with 0.5-ms onset and offset ramps). The spectra of the Gaussian noise bursts were bandpassed between 1 and 30 kHz with abrupt cutoffs. The stimulus levels were 20 to 40 dB above the unit's threshold in 5-dB steps. A total of 40 trials was delivered for each combination of stimulus location and stimulus level; locations and levels were varied in a pseudorandom order.

Spectrally-filtered noise, consisting of 80-ms bursts of narrowband noise, notched noise, and highpass noise, were always presented at 80 or 100° elevation. We chose those locations to present the spectrally-filtered noise because cats' head-related transfer functions typically were flattest for these locations. The narrowband noise had a flat center 1/6-oct wide and skirts that fell off at 128 dB per octave. The center frequencies (F_c 's) of the narrowband noise stimuli that we used were usually from 4 to 18 kHz in 1-kHz steps. In some cases, the range of F_c 's were extended to 28 kHz. The reject bands for the notch stimuli had a flat center 1/6-oct, 1/2-oct, or 1-oct wide and skirts that rose at 128 dB per octave. The depth of the notch was 40 dB and the widths at the top were 0.792, 1.125, or 1.625 octave. The F_c 's of the notch typically ranged from 4 to 18 kHz in 1-kHz steps. The highpass noise had a positive slope of 128 dB per octave. The 3-dB cutoff frequencies of the highpass noise ranged from 6 to 20 kHz in 1-kHz steps. The sound levels of the spectrally-filtered noise were equalized by root-mean-squared power.

Perceptually, two sounds of equal root-mean-squared power that differ in spectral shape might produce different loudnesses. Therefore, the stimulus levels all were expressed as stimulus levels above unit's threshold for each type of spectrally-filtered noise. Stimulus levels 20, 30, and 40 dB above threshold were used for the spectrally-filtered stimuli. A total of 20 trials was delivered for each combination of stimulus F_c or cutoff frequency and stimulus level; frequencies and levels were varied in a pseudorandom order.

Narrowband stimuli at 1 - 3 F_c 's also were varied across a range of elevations to study the elevation sensitivities of neurons to the narrowband noise. The narrowband noise of selected F_c 's were presented from the 14 loudspeakers in the vertical plane, one loudspeaker at a time. The stimulus levels for each F_c were 20, 30, and 40 dB above threshold. A total of 20 trials was delivered for each combination of stimulus location and stimulus level; locations and levels were varied in a pseudorandom order.

Measurement of head-related transfer functions (HRTFs) of the external ears was carried out in all cats after the physiological experiments. A 1/2" probe microphone (Larson-Davis model 2540) was inserted into the ear canal through a surgical opening at the posterior base of the pinna. The probe stimuli delivered from each of the 14 loudspeakers in the median plane were pairs of Golay codes (Golay 1961; Zhou et al. 1992) that were 81.92 ms in duration. Recordings from the microphone were amplified and then digitized at a rate of 100 kHz, yielding a spectral resolution of 12.2 Hz from 0 to 50 kHz. We divided from the amplitude spectra a common term that was formed by the root-mean-squared sound pressure averaged across all elevations. Removal of the common term left the component of each spectrum that was specific to each location; we have referred to that term previously as the directional transfer function (Middlebrooks

and Green 1990), but now adopt the term HRTF in agreement with common usage. We convolved each HRTF in the linear frequency scale with a bank of bandpass filters to transfer it to a logarithmic (i.e., octave) scale (Middlebrooks 1999a). The filter bank consisted of 118 triangular filters. The 3-dB bandwidth of the filters was 0.0571 octave, filter slopes were 105 dB per octave, and the center frequencies were spaced in equal intervals of 0.0286 octave from 3 to 30 kHz yielding 118 bands. The interval of 0.0286 was chosen to give intervals of 2% in frequency.

Data Analysis

The goals of the data analysis were, first, to map the correspondence of broadband sound-source elevations with cortical spike patterns and, then, to associate spike patterns elicited by various filtered sounds with broadband source elevations. Artificial neural networks were employed to map spike patterns onto source elevations. Networks were constructed using MATLAB Neural Network Toolbox (The Mathworks, Natick, MA) and were trained with the back-propagation algorithm (Rumelhart et al. 1986). The architecture, as detailed in Middlebrooks et al. (1998), consisted of a 4-unit hidden layer with sigmoid transfer functions and a 2-unit linear output layer. The inputs to the neural network were spike density functions expressed in 1-ms time bins. The spike density functions were derived from a bootstrap averaging procedure (Efron and Tibshirani 1991) in which each spike density function was formed by repeatedly drawing 8 samples with replacement from the neural responses to a particular stimulus condition. The two output units of the neural network produced the sine and cosine of the stimulus elevation, and the arctangent of the two outputs gave a continuously varying output in degree in elevation, i.e., the polar angle around the interaural axis. We did not constrain

the output of the network to any particular range, so the scatter in network estimation of elevation sometimes fell outside the range of locations to which the network was trained (i.e., from -60 to $+200^\circ$). Typically, we formed 20 bootstrapped training patterns from the odd-numbered trials of the neural responses to the broadband noise stimuli and used them to train the artificial neural network. The trained network was then subjected to testing with patterns consisted of 100 bootstrapped trials derived from either the even-numbered trials of the neural responses to broadband noise or the entire set of neural responses to spectrally-filtered noise.

Results

Usable unit and unit-cluster data were obtained at 389 recording sites in 33 multichannel probe placements in auditory area A2 in 8 cats. All of the A2 units showed relatively broad frequency tuning that was defined by frequency tuning curves that were at least one octave wide at 40 dB above threshold. For 60.2% of the units, the tuning curve of each unit spanned the entire mid-frequency range of 6 - 19 kHz. In the following, we report the general properties of these units in response to broadband and narrowband noise stimulation at various source elevations. We then examine the sensitivity of units for the elevation of broadband noise sources. A quantitative model that predicts human judgements of the locations of narrowband sounds is adapted for the cat, then model predictions are compared with the locations signaled by cortical neurons in response to narrowband stimuli. The neural responses to notch stimuli are also analyzed using the neural-network algorithm. Next, we compare the elevation sensitivity of the neural responses to highpass noise stimulation with that of neural responses to

narrowband noise stimulation. Finally, we examine the consequences for localization coding of excluding information conveyed by the timing of spikes.

General Properties of Neural Responses to Broadband and Narrowband Stimuli

As we demonstrated in the previous study (Xu et al. 1998), A2 units showed broad elevation tuning in response to broadband noise stimulation. An example of the spike patterns of one representative unit (9806C02) in response to broadband noise is represented by a raster plot in Figure 4.1A. Sound-source elevation is plotted on the ordinate and the post-stimulus onset time is plotted on the abscissa. Each dot represents one spike recorded from the unit. Only 20 trials of responses for each stimulus condition elicited at 30 dB above threshold are shown here. One can see subtle changes in the numbers and distribution of spikes and in the latencies of the spike patterns from one elevation to another. The elevation tuning of the unit's mean spike counts in response to broadband noise at 20 to 40 dB above threshold in 5-dB steps is plotted in Figure 4.1D. Spike counts showed some elevation tuning at the lowest stimulus level but tuning flattened out at higher stimulus levels. We quantified the elevation tuning of spike counts by the average modulation of the spike counts by sound-source elevation across 20, 30, and 40 dB above threshold. The modulation for the unit in Figure 4.1A, averaged across sound levels, was 59.2%. Across the whole population of 389 units that we studied using broadband noise, the median of the average modulation was 47.8%, which was comparable with our previous report (Xu et al. 1998).

Narrowband stimuli produced weaker elevation tuning than did broadband stimuli. The raster plots (Figure 4.1, B and C) show the spike patterns of the same unit elicited by narrowband noise centered at F_c of 6 and 16 kHz, respectively. Spike

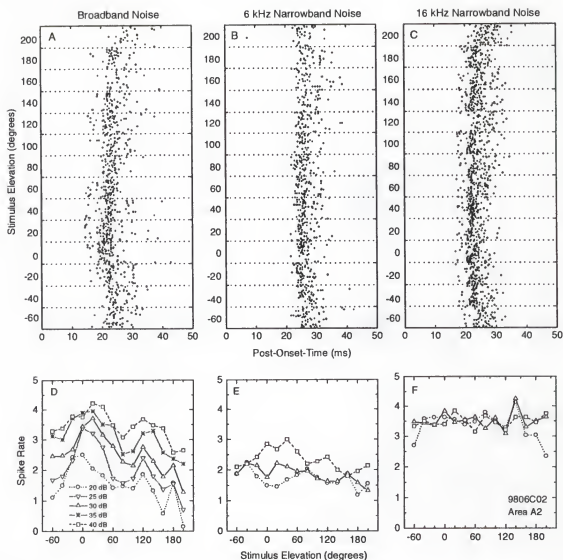


Figure 4.1. Unit responses elicited by broadband and narrowband noise (unit 9806C02). A: Raster plot of responses to broadband sounds presented from 14 locations in the median plane. Each dot represents one spike from the unit. Each row of dots represents the spike pattern recorded from one presentation of the stimulus at the location in elevation indicated along the vertical axis. Only 20 trials recorded at each elevation are plotted. Stimuli were 80 ms in duration and 30 dB above threshold. B and C: Raster plots of responses to 1/6-oct narrowband noise with center frequencies at 6 and 16 kHz, respectively. Other conventions are the same as in A. D: Spike-rate-versus-elevation profiles for the responses to broadband stimulation. Each line represents the spike-rate-versus-elevation profile at one of the five stimulus levels (i.e., 20, 25, 30, 35, and 40 dB above threshold). E and F: Spike-rate-versus-elevation profiles for the responses to 6- and 16-kHz narrowband stimulation, respectively. Stimulus levels were 20, 30, and 40 dB above threshold. Symbols and line types match those in D that represent the equivalent levels.

patterns showed less variation from one elevation to another than did those elicited by broadband stimuli. On the other hand, spike patterns showed considerable variation across F_c . Fewer spike counts were elicited by 6-kHz narrowband noise than by 16-kHz narrowband noise. The spike patterns elicited by 16-kHz narrowband noise usually started with a single short-latency (< 20 ms) spike followed by a silent period of about 3 ms and then several spikes at short interspike intervals (Figure 4.1C). These firing patterns resembled those elicited by broadband noise at $+20$ to $+60^\circ$ elevation (Figure 4.1A). Figure 4.1, E and F, plots the elevation tuning of the unit in response to the two narrowband stimuli at 20, 30 and 40 dB above threshold. The elevation tuning curves were flatter than those of broadband noise stimulation; the average modulation of elevation was 30.6 and 20.8% for 6- and 16-kHz narrowband stimulation, respectively. Across the sample of 158 units that we recorded using narrowband stimuli, the median of the average modulation of spike counts by elevation of narrowband noise was 39.9%.

Network classification of responses to broadband stimulation

Results from artificial-neural-network analysis of the spike patterns elicited by broadband noise stimulation were comparable with our previous report (Xu et al. 1998). The A2 neurons could code sound-source elevation with their spike patterns with various degree of accuracy. As an example, the network analysis of the spike patterns of the same unit as in Figure 4.1A elicited at 30 dB above threshold is shown in Figure 4.2A. Each plus (+) represents the network estimate of elevation based on one spike pattern, and the solid line indicates the median direction of responses at each stimulus source elevation. In general, the neural-network estimates scattered around the perfect performance line (---). Some large deviations from the targets were seen at certain

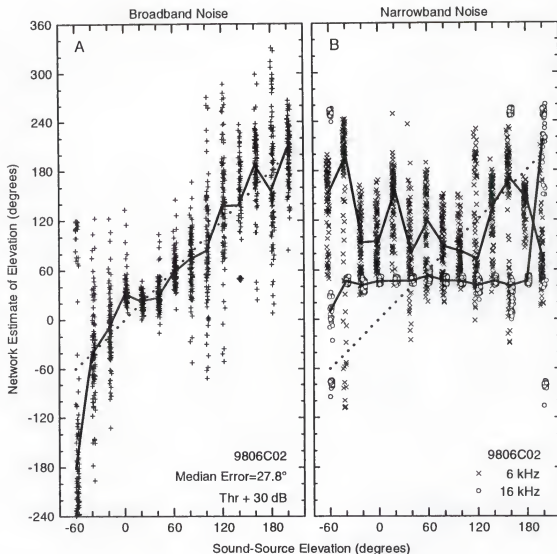


Figure 4.2. Network analysis of spike patterns of the same unit (9806C02) as in Figure 4.1. A: Network performance in classifying spike patterns elicited by broadband noise at 30 dB above threshold. Each symbol represents the network output in response to input of one bootstrapped patterns. The abscissa represents the actual stimulus elevation, and the ordinate represents the network estimate of elevation. The solid line connects the median directions of network estimates for each stimulus location. Perfect performance is represented by the dashed diagonal line. B: Network classification of spike patterns elicited by narrowband noise of center frequencies at 6 kHz (o) and 16 kHz (x). The neural network was trained with spike patterns elicited by broadband noise at 5 roving levels (20, 25, 30, 35, and 40 dB above threshold) and was tested with those elicited by narrowband noise at 30 dB above threshold. Other conventions are the same as in A.

locations in elevation (e.g., -60° in this example). We calculated the median error of the neural-network estimates as a global measure of network performance. The neural network classification of the spike patterns of the unit shown in Figure 4.2A yielded a median error of 27.8° , which was among the smallest in our sample of recordings with broadband noise stimuli.

Across all the 389 units that we studied with broadband noise stimuli, the median errors of the network performance averaged 41.7 and 50.4° for stimulus levels of 20 and 40 dB above threshold, respectively, ranging from 19.9 to 67.2° . The averaged median errors were 3 to 4° larger than in the data set that we reported previously (Xu et al. 1998). This small difference probably was due to differences in unit recording and spike sorting techniques. Nonetheless, the bulk of the distribution of median errors was substantially better than chance performance of 65° . The distribution of the median errors was unimodal. We selected the half of the distribution with the lowest median errors at 40 dB above threshold (194 units; median errors $< 50.4^\circ$) for analysis of responses to filtered sounds. Among those 194 elevation-sensitive units, 73 units were tested using narrowband noise of fixed F_c 's at various elevations. Using stimuli fixed in elevation at $+80$ or $+100^\circ$, all 194 elevation-sensitive units were tested with narrowband noise of varying F_c 's, 127 were tested with notches of varying F_c 's and 74 were tested using highpass noise stimuli.

Neural Network Classification of Responses to Narrowband Stimulation

The spike patterns of narrowband noise stimulation presented from 14 midline elevations showed less variation across locations than did spike patterns to broadband noise stimulation, as shown in Figure 4.1. When we trained the artificial neural network

with spike patterns elicited by broadband stimulation and used this trained network to classify the spike patterns elicited by narrowband stimulation, we found that the network outputs tended to cluster around certain locations in elevation, regardless of the actual source locations. Figure 4.2B shows an example of the neural-network outputs for one of the elevation-sensitive units (9806C02); the spike patterns of this unit are plotted in Figure 4.1, B and C. The network estimates of elevation for 6-kHz narrowband noise are plotted with crosses (x) and those for 16-kHz narrowband noise are plotted with circles (o). The neural-network outputs for spike patterns elicited by the 6-kHz narrowband noise tended to scatter in the upper-rear quadrant, whereas those for spike patterns elicited by 16-kHz narrowband noise tended to point around 50° above the front horizon. The network estimates of elevation for the neuronal responses to narrowband stimulation were dependent on the center frequency but independent of the actual source location.

In the following analysis, we tested the neural responses to narrowband stimulation of different F_c 's presented at a fixed location. In this test, we trained the neural network with spike patterns elicited by broadband noise at 5 roving levels (20, 25, 30, 35, and 40 dB above threshold). After the neural network learned to recognize the spike patterns of broadband stimulation according to sound-source elevation, the trained network was used to classify the neural responses to narrowband noise stimulation of varying F_c 's.

An example of the spike patterns elicited by broadband noise and narrowband noise from one of our elevation-sensitive units (9806C16) is shown in Figure 4.3 in a similar format to that of Figure 4.1. Broadband noise stimuli were presented from 14

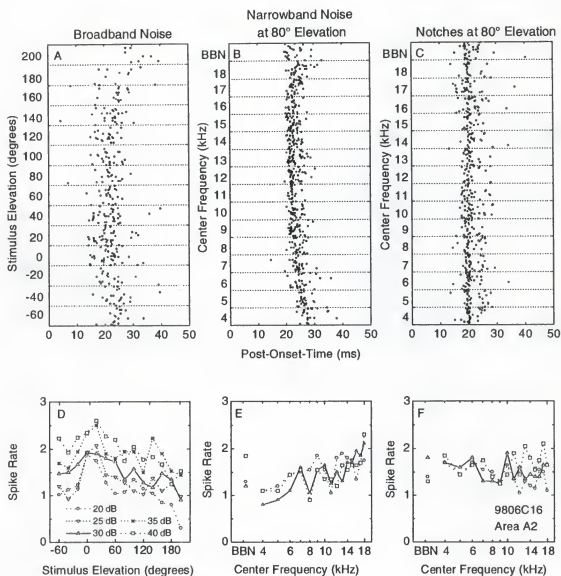


Figure 4.3. Unit responses elicited by broadband, narrowband, and notched noise (unit 9806C16). A: Raster plot of responses to broadband stimulation presented from 14 locations in the median plane. Conventions as Figure 4.1A. B: Raster plots of responses to narrowband noise of various center frequencies. The narrowband stimuli were presented from +80° elevation. The narrowband center frequencies were from 4 to 18 kHz as indicated along the vertical axis with BBN indicating spike patterns elicited by broadband sounds presented at +80° elevation. Stimuli were 20 dB above threshold. C: Raster plots of responses to 1/6-oct notched noise of center frequencies ranging from 4 to 18 kHz in 1-kHz steps. Other conventions are the same as in B. D: Spike-rate-versus-elevation profiles for the responses to broadband stimulation. Conventions as Figure 4.1A. E and F: Spike-rate-versus-center-frequency profiles for the responses to narrowband and notched noise, respectively. Stimulus levels were 20, 30, and 40 dB above threshold. Symbols and line types match those in D that represent the equivalent levels. BBN on the abscissa indicates spike rate elicited by broadband noise.

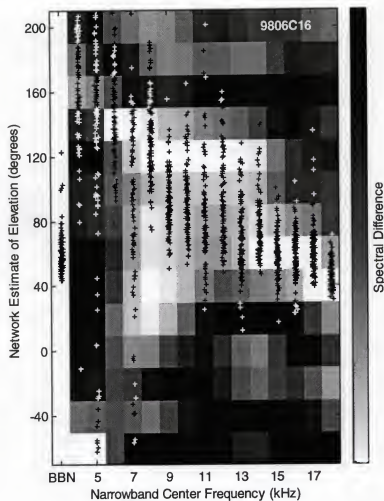


Figure 4.4. Network estimates of elevation. The network analysis was based on the responses to narrowband sounds that varied in center frequency; the neural responses of the unit (9806C16) are shown in Figure 4.3. The neural network was trained with spike patterns elicited by broadband noise presented from 14 elevations at 5 roving levels (20, 25, 30, 35, and 40 dB above threshold) and was tested with those elicited by narrowband noise at 30 dB above threshold. Each column of symbols represents network outputs for spike patterns elicited by narrowband noise of a given center frequency as indicated along the abscissa. BBN indicates the network responses to spike patterns elicited by broadband noise. All stimuli were presented from $+80^\circ$ elevation. The background of gray-scale rectangles for the narrowband stimuli represents the acoustical model predictions that are based on the spectral differences between the narrowband stimulus spectra and the head-related transfer functions at each elevation. Values of the spectral differences were scaled to span the full lightness between the extremes of black and white. White and light gray indicate small spectral differences and the network estimates that fall in those regions are plotted in black. Black and dark gray indicate large spectral differences and the network estimates that fall in those regions are plotted in white.

elevations (Figure 4.3, A). The narrowband stimuli of F_c 's from 4 to 18 kHz in 1-kHz steps were presented at +80° elevation (Figure 4.3, B). Only 20 response patterns in each stimulus condition are shown here. The spike rate tuning of the unit at 5 different stimulus levels of broadband noise and 3 different stimulus levels of narrowband noise are plotted in Figure 4.3, D and E. Both elevation tuning of the broadband noise and the frequency tuning to narrowband noise were fairly broad.

Figure 4.4 shows the network estimate of elevation based on responses of the same unit (9806C16) to narrowband sounds that varied in F_c . Each column of plus signs represents the network output for one F_c . The background of gray-scale rectangles represents the acoustical model that is described in the next section. In this case, the network estimates of elevations for the narrowband noise data tended to shift monotonically to lower elevations as F_c 's increased. The network outputs for broadband noise data are shown on the stripe of white background. The median direction of the network estimation for the broadband noise data was +59.9°, which was about 20° off the location (+80° elevation) from which the broadband noise was actually presented.

Figure 4.5 shows an example from a unit (9803A02) in a different cat. Narrowband noise stimuli with 10 different F_c 's (7 to 16 kHz in 1-kHz steps) were presented at +80° elevation. In this case, the network estimates of elevation varied somewhat erratically with F_c of the stimuli. The median direction of the network estimation for the broadband noise data was +93.7°, which was 13.7° off the target (+80° elevation) where the broadband noise was actually presented.

The Model of Spectral Shape Recognition

In a previous human psychophysical study, we presented a quantitative model

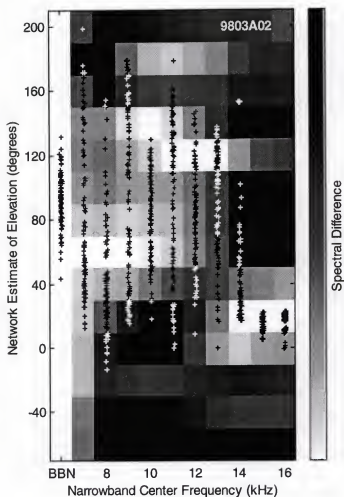


Figure 4.5. Network analysis of spike patterns and model predictions in response to narrowband stimulation. This example is taken from a unit (9803A02) in a different cat from that shown in Figure 4.4. Narrowband center frequencies varied from 7 to 16 kHz in 1-kHz steps. Other conventions are the same as in Figure 4.4.

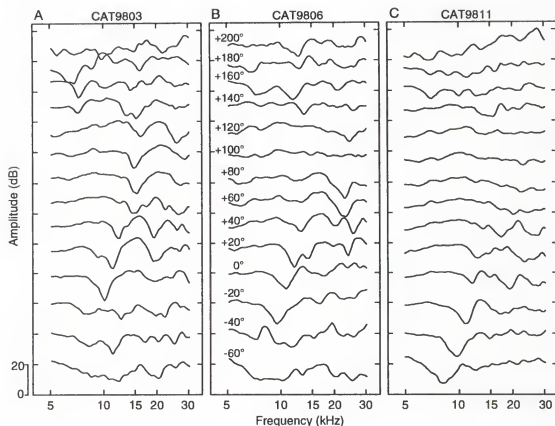


Figure 4.6. Head-related transfer functions (HRTFs) in the median plane measured from left ears of 3 cats. The measurements and process of HRTFs are described in detail in METHODS. Starting from the bottom, each line represents a HRTF for one of the 14 midline elevations from -60 to $+200^\circ$, as indication on the left in B. A: cat9803. B: cat9806. C: cat9811.

that used a comparison of stimulus spectra with head-related transfer functions (HRTFs) to predict listeners' judgements of the locations of narrowband sounds (Middlebrooks 1992). In the present study, we adapted that model to the cat as a means of simulating cats' location judgements. The model was adapted by substituting feline HRTFs for human HRTFs and by extending the frequency range of the analysis to higher frequencies to accommodate the cats' higher audible range.

Figure 4.6 shows examples of HRTFs for all the 14 midline elevations measured in the left ears of 3 cats (A, cat9803; B, cat9806; C, cat9811). There were considerable individual differences among cats. In general, however, spectral features, such as peaks and notches, tended to increase in center frequency as sound sources increased in elevation in the front (-60 to $+80^\circ$) and, to a lesser degree, in the rear ($+200$ to $+100^\circ$). The most systematic variation occurred in the mid-frequency region (5 - 18 kHz), which has been emphasized in previous studies of the cat HRTFs (Musicant et al. 1990; Rice et al 1992). In most cats, HRTFs at overhead locations ($+80$ to $+100^\circ$ elevation) were relatively flat, although exceptions did occur (e.g., Figure 4.6A). Differences in the midline HRTFs measured from the left and right ears of a given cat tended to be smaller than the differences among cats. The median spectral differences between left and right ears across all 8 cats was 10.4 dB^2 , whereas the median spectral differences between left ears of all 28 pairs of cats was 14.5 dB^2 . In the spectral recognition model that predicted the narrowband noise localization behavior of the individual cats, we used the HRTFs measured from each cat's own left ear, i.e., contralateral to the physiological recording site.

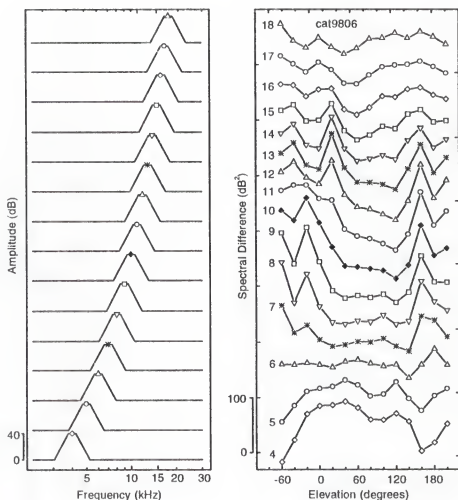


Figure 4.7. Spectral differences between the narrowband stimulus spectra and HRTFs. Left panel: Spectra of narrowband noise of center frequencies from 4 to 18 kHz in 1-kHz steps. Symbols represent the center frequencies. Right panel: Spectral differences. Each line represents the spectral differences between the spectrum of the narrowband noise of a given center frequency as indicated on the left of the line and the HRTFs measured from 14 elevations as indicated by the abscissa. HRTFs were taken from cat9806 (Figure 4.6, B).

We defined a metric to quantify the similarity between the narrowband noise stimuli and the HRTFs. First, the stimulus spectrum was added to the HRTFs of the elevation at which the stimulus was presented. Next, we subtracted, frequency by frequency, the log-magnitude spectrum of each HRTF from that of each narrowband stimulus. Then, we computed the variance of each difference distribution across all frequencies. We referred to the variance of the difference distribution as the *spectral difference*. The smaller the spectral difference, the more similar are the stimulus spectrum and the HRTF. Figure 4.7 illustrates how this computation was accomplished for the data from one of the cats (cat9806). The amplitude spectra of the 1/6-oct narrowband noise stimuli with F_c 's from 4 to 18 kHz in 1-kHz steps are shown in the left panel of Figure 4.7. The right panel of Figure 4.7 plots the spectral differences. The abscissa in the right panel of Figure 4.7 represents the source elevations at which the 14 HRTFs were measured; those HRTFs are shown in Figure 4.6B. Each line in the right panel of Figure 4.7 represents the spectral difference between one narrowband noise stimulus (Figure 4.7, left panel) and the 14 HRTFs (Figure 4.6B). The symbols used for the lines match the symbols used to represent the F_c 's of the narrowband noise spectra shown in the left panel of Figure 4.7.

Our model predicts that an individual animal's judgement of a narrowband sound source would be biased towards elevations at which the spectral differences are small. If the responses of cortical neurons are influenced by the narrowband noise stimulus in the same way as is the behavior of the animal, the spike patterns elicited by narrowband noise of a particular F_c should resemble the spike patterns elicited by broadband noise at source elevations at which the spectral differences are small. In terms of the artificial-neural-

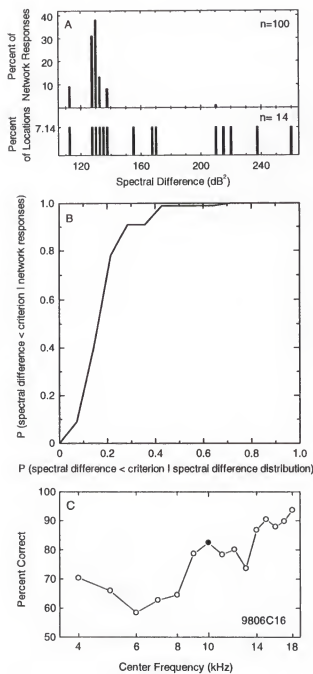
network algorithm, a neural network trained with spike patterns of broadband noise stimulation should localize the spike patterns of narrowband noise stimulation to locations in which small spectral differences are found.

Figures 4.4 and 4.5 show the output of the acoustical model in register with the network estimates of elevation based on neural responses to narrowband stimuli. For each narrowband F_c , values of the spectral differences were scaled to span the full lightness between the extremes of black and white. White and light gray indicate small spectral differences and the network estimates that fall in those regions are plotted in black. Black and dark gray indicate large spectral differences and the network estimates that fall in those regions are plotted in white. In both figures, neural network outputs tend to fall within white-to-light-gray areas on the background, i.e., regions with small spectral differences. Inter-cat differences in HRTFs resulted in individual differences in spectral differences, as indicated by differences between Figures 4.4 and 4.5 in the background patterns. The elevation estimates based on physiological data also showed individual differences, which presumably resulted in part from differences in the HRTFs that shaped the input to the neurons.

Correspondence of Physiology with Behavioral Simulation

The neural-network analysis of the spike patterns elicited by narrowband noise stimuli had a distinct distribution for each F_c . By our hypothesis, the distribution was more likely to be concentrated in the location at which the spectral differences were small. We tested this model against the alternative hypothesis that the distribution of the network estimates across locations is random. The test was adapted from one used in our previous psychophysical study (Middlebrooks 1992), which was in turn adapted from

Figure 4.8. Correspondence between model prediction and network outputs. Data are from the example shown in Figure 4.4 (unit 9806C16). A: Distribution of spectral differences. The lower panel represents the distribution of the spectral differences between 10-kHz narrowband noise and the 14 HRTFs. Data are taken from the seventh line from the bottom in Figure 4.7. The upper panel represents the distribution of the spectral difference at the elevations corresponding to the network estimates. Data are from the network estimates of elevation for 10-kHz narrowband noise (eighth column from left in Figure 4.4). B: Receiver-operating-characteristic (ROC) curve. Data are derived from A. We varied a criterion from left to right on the abscissa of A and plotted in B the percentages of two distributions in A that fell below the criterion. The area under the ROC curve, 0.825 in this case, represents the fraction of physiological trials in which the network estimate fell at an elevation at which the spectral difference was smaller than the median spectral difference across all elevations. If the network outputs were random, the ROC curve would be close to the main diagonal line and the area under it would be 0.50. The area under the ROC curve is referred to as percent correct thereafter. C: Percent correct for unit 9806C16. We calculated and plotted the percent correct associated with the 15 different narrowband center frequencies (abscissa) that we tested for this unit. The filled circle at 10 kHz represents the data that are derived from A and B.



Signal Detection Theory (Green and Swets 1966). The procedure is demonstrated in Figure 4.8, using the 10 kHz data shown in Figure 4.4. We first plotted in the lower panel of Figure 4.8A the distribution of the spectral differences calculated from the spectrum of 10-kHz narrowband noise and the 14 HRTFs. We then plotted in the upper panel of Figure 4.8A the distribution of the spectral difference at the elevations corresponding to the network estimates. Network estimates clustered at locations in which the spectral differences were relatively small. Next, we varied a criterion from left to right on the abscissa of Figure 4.8A and plotted in Figure 4.8B the percentages of distributions in Figure 4.8A that fell below the criterion; this formed a receiver-operating-characteristic (ROC) curve. The area under the ROC curve represents the fraction of physiological trials in which the network estimate fell at an elevation at which the spectral difference was smaller than the median spectral difference across all elevations. If the network outputs were random, the ROC curve would be close to the main diagonal line and the area under it would be .50. In this particular example, the area under the ROC curve was .825, or 82.5% correct. In Figure 4.8C, we plotted the percent correct associated with the 15 different narrowband noise F_c 's that we tested for this unit. Note that all values of percent correct were larger than chance performance of 50%. The filled circle at 10 kHz represents the data that were derived from Figure 4.8, A and B.

Figure 4.9 shows the distribution of percent correct for all the narrowband F_c 's that we used across the 194 elevation-sensitive units. The abscissa represents the narrowband noise F_c 's. The solid line and two dashed lines represent the median, the upper and the lower quartiles of the distribution of percent correct, respectively. The dotted line represents the prediction of 50% based on chance performance. The number

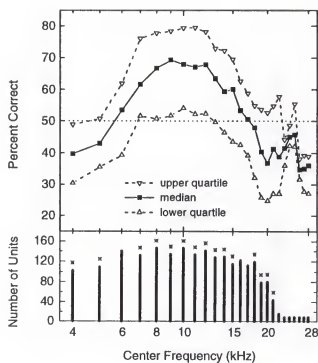


Figure 4.9. Distribution of percent correct for all narrowband center frequencies across the sample of units. The narrowband center frequency is represented by the abscissa. The solid line and two dashed lines represent the median, the upper and the lower quartiles of the distribution of percent correct, respectively. The dotted line represents the chance performance of 50%. The number of units that we tested with narrowband noise of each center frequency is indicated by the bars in the lower panel. The asterisks over the bars indicate the center frequencies at which percent correct values statistically significant from 50% (two-tailed t test, $P < 0.05$).

of units that we tested with narrowband noise of each F_c is shown by the bars in the lower panel of Figure 4.9. The asterisks over the bars indicate F_c 's at which percent correct values statistically significant from 50% (two-tailed t test, $P < 0.05$). The majority of our units had a percent correct $>50\%$ in the frequency range between 7 and 15 kHz. That indicates that the model prediction and the neural responses correspond well with each other in that mid-frequency range. On the other hand, the distribution of percent correct at very low frequency (4 and 5 kHz) as well as at high frequency (>17 kHz) was below the chance performance line of 50%, which suggested that the model poorly predicted the neural responses at those frequency ranges. The poor performance at low frequencies presumably reflects the fact that most units in A2 respond weakly if at all to low frequency sounds (Xu et al. 1998). Also, the HRTFs recorded from the eight cats used in this study generally did not show direction-dependent changes in spectral features at frequency < 6 or 7 kHz. Consistent with other reports (Musicant et al. 1990; Rice et al. 1992), we found that the high-frequency region (>17 kHz) in the HRTFs was highly complex and irregular (Figure 4.6, for example). As we consider in the Discussion, cats show accurate localization when stimulus spectra are limited to the mid-frequency region but not when limited to high or low frequencies (Huang and May 1996a).

Neural Responses to Stimuli Containing a Narrowband Notch

Spectral notches are among the most prominent features in the HRTFs. Several authors have suggested that a single spectral notch in each ear could uniquely specify the source elevation in the median plane (Musicant et al. 1990; Neti et al. 1992; Rice et al. 1992). For that reason, one might predict that a notch in the source spectrum would

signal an erroneous vertical location. In this section, we tested such a hypothesis using notched noise stimuli.

Spike patterns elicited by notch stimuli generally were more homogeneous than those elicited by narrowband noise. An example of the neural responses to 1/6-oct notch stimuli is shown in Figure 4.3C. Data were obtained from the same unit as in Figure 4.3, A and B. The spike patterns varied somewhat less prominently as a function of the notch F_c 's, compared to those elicited by bandpass stimuli. The spike-count tuning to notches was only weakly modulated by the notch F_c 's, as shown in Figure 4.3F.

Using neural networks that were trained with spike patterns elicited by broadband noise, we evaluated the elevation coded by the spike patterns elicited by the notches. Generally, neural network outputs showed little variation with varying notch F_c 's. Figure 4.10 plots the network estimates of elevation for the spike patterns of the unit shown in Figure 4.3C. For F_c 's < 12 kHz, the network output for the notches did not differ from those for broadband noise. Some variation of the estimated elevation was seen for F_c 's > 12 kHz. However, the network estimated elevation did not follow the predictions made by matching the F_c 's of stimulus notches with the notches in the HRTFs. For example, a 10-kHz notch matched best with the notch in the HRTF measured from -20° elevation (Figure 4.6B), yet the network outputs for this F_c were clustered between 0 and 80° elevation. A 13-kHz notch stimulus matched with the notches in the HRTFs measured from $+40$, $+140$, $+180$, and $+200^\circ$ elevation (Figure 4.6B). The network outputs for that F_c were mostly concentrated between $+40$ and $+130^\circ$ elevation. Therefore, the variation shown in the spike patterns and network outputs for the notch stimulation was probably more complicated than can be explained by a single-notch matching scheme. Our

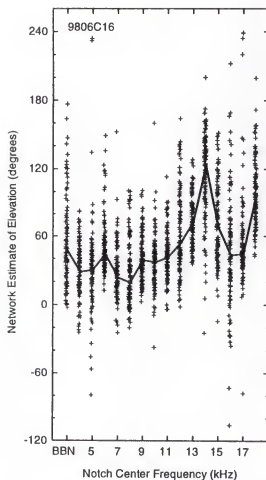


Figure 4.10. Network analysis of spike patterns elicited by notched noise. Spike patterns of the unit (9806C16) elicited by notches are shown in Figure 4.3C. The neural network was trained with spike patterns elicited by broadband noise presented from 14 elevations at 5 roving levels (20, 25, 30, 35, and 40 dB above threshold) and was tested with those elicited by notched noise at 30 dB above threshold. Each symbol represents a network estimate of elevation for 1 bootstrapped pattern. All stimuli were presented from $+80^\circ$ elevation. Notch filter center frequencies were from 4 to 18 kHz in 1-kHz steps. BBN indicates the network responses to spike patterns elicited by broadband noise.

systematic analysis of the data from the population of 127 units recorded using spectral notches of various widths (1/6, 1/2, or 1 octave) produced results that were inconsistent with the single-notch matching hypothesis.

Comparison of Narrowband Noise Results to Highpass Noise Data

We considered two alternative hypotheses that might account for the variation in unit spike patterns in response to varying F_c of narrowband sounds. The first was that the magnitude of unit responses simply reflected the amount of overlap between the narrowband stimulus spectrum and the units' frequency response area. The alternative was that units were sensitive to the frequencies of specific elements of spectral shape such as spectral slopes or changes in slope. We attempted to differentiate between these hypotheses by testing unit responses to stimuli that differed markedly in frequency content but that shared a spectral feature. Specifically, we compared responses to narrowband sounds with highpass noise. This test was motivated by recent psychophysical results from our laboratory showing that human listeners tend to make similar elevation judgments when the low frequency cutoffs of narrowband and highpass stimuli are equal (Macpherson and Middlebrooks 1999).

An example of the spike patterns of one of the units (9811C03) in response to broadband, narrowband, and highpass noise is shown in Figure 4.11, A, B and C, respectively. The ordinates of Figure 4.11, B and C, represent narrowband F_c 's and highpass cutoff frequencies. Only 20 trials of responses for each stimulus condition elicited at 30 dB above threshold are plotted here. The elevation tuning of the unit spike counts in response to broadband noise at various sound levels is plotted in Figure 4.11D.

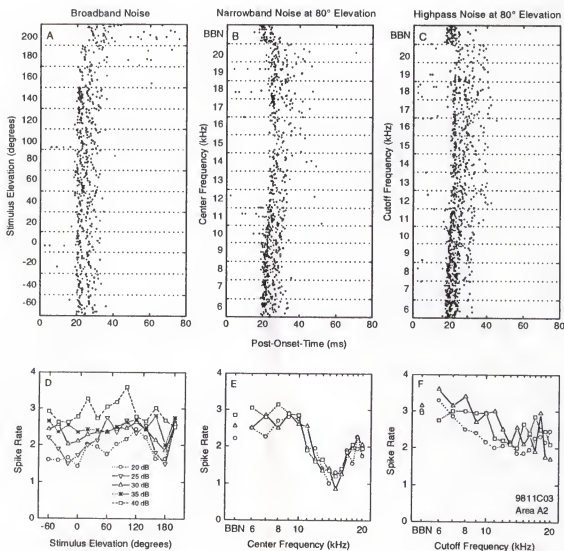


Figure 4.11. Unit responses elicited by broadband, narrowband, and highpass noise (unit 9811C03). C and F plot responses elicited by highpass noise of cutoff frequencies from 6 to 20 kHz in 1-kHz steps. Other conventions are the same as in Figure 4.3.

The distribution of spikes in time (Figure 4.11A) varied with source location whereas spike-count tuning (Figure 4.11D) was fairly broad. The tuning of spike counts to narrowband noise F_c 's and highpass noise cutoff frequencies is shown in Fig 11, E and F. The variations in spike counts for the two types of noise were quite different, whereas their temporal patterns (Figure 4.11, B and C) were rather similar.

Following the procedure that we used for unit responses to narrowband noise, we used neural network to obtain estimates of elevation based on unit responses to highpass noise. We trained the neural network with spike patterns elicited by broadband noise at 5 levels (20, 25, 30, 35, and 40 dB above threshold) then used network to classify the spike patterns elicited by narrowband and highpass noise stimulation of various frequency contents. Figure 4.12 shows network outputs based on the spike patterns shown in Figure 4.11. Narrowband and highpass filter functions are shown in the upper panel; network outputs are shown in the lower panel. Filled triangles represent network outputs for spike patterns elicited by narrowband stimuli and open triangles represent those for spike patterns elicited by highpass stimuli. The narrowband noise F_c 's are indicated on the upper abscissa and the highpass cutoff frequencies on the lower abscissa. The narrowband F_c 's are one kHz above the highpass cutoff frequencies. The reason for such an alignment of highpass cutoff frequencies and narrowband F_c 's is that it provides an approximate match for the positive slopes (i.e., lower cutoffs) of the spectra of the two types of noise stimuli across the frequency range that we used (Figure 4.12, upper panel). The amplitude spectra in the upper panel of Figure 4.12 align with the network outputs for the same stimuli in the lower panel. The network estimated elevation varied as a function of highpass cutoff frequencies and narrowband F_c 's. The network elevation

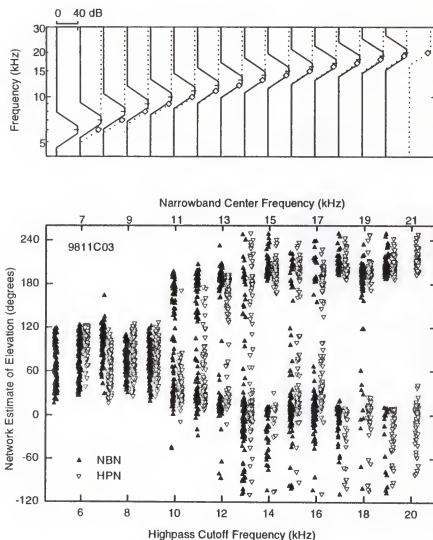


Figure 4.12. Comparison of network classification of the spike patterns elicited by narrowband and highpass noise. Upper panel: Spectra of narrowband and highpass stimuli are plotted by solid and dotted lines, respectively. The narrowband center frequencies are represented by short lines (-) and the highpass cutoff frequencies are represented by open diamonds (\diamond). The narrowband center frequencies are one kHz above the highpass cutoff frequencies, which provides an approximate match for the positive slopes of the spectra of the two types of noise stimuli. Lower panel: Open and filled triangles represent the network outputs for spike patterns elicited by narrowband and highpass noise, respectively. The neural responses of the unit (9811C03) are shown in Figure 4.11. The neural network was trained with spike patterns elicited by broadband noise presented from 14 elevations at 5 roving levels (20, 25, 30, 35, and 40 dB above threshold) and was tested with those elicited by narrowband or highpass noise at 30 dB above threshold. The narrowband center frequencies indicated on the upper abscissa are one kHz above the highpass cutoff frequencies indicated on the lower abscissa.

estimates for the spike patterns elicited by both types of noise stimuli were very similar when the positive slopes of the spectra of the highpass noise matched those of the narrowband noise.

The network elevation estimates based on response to highpass stimuli could be explained qualitatively by comparing stimulus spectra with the individual HRTFs. The unit shown in Figure 4.12 was recorded from cat9811 whose HRTFs are plotted in Figure 4.6C. The network outputs for the highpass data formed three patterns depending on cutoff frequencies. First, for cutoffs < 9 kHz, the majority of network estimates fell between $+60$ and $+120^\circ$ elevation. When cutoffs were < 9 kHz, flat pass bands extended across most of the mid- and high-frequency regions, thus providing valid spectral cues to the actual source location of 80° . Also, HRTFs from those high elevations tended to be relatively flat (Figure 4.6C). Second, for cutoffs between 9 and 13 kHz, the network outputs showed a transition from a cluster at one location to two separate clusters. Highpass noise of cutoffs between 9 and 13 kHz had positive slopes that mimicked the positive slopes in the HRTFs from lower elevations from -60 to $+20^\circ$. The network outputs tended to favor locations slightly higher than those locations. Such biases were noticed in our previous report that for sound sources at lower elevations the network estimates tended to point above the source locations (Xu et al. 1998). Thirdly, for cutoffs > 13 kHz, the network estimates pointed to two regions in elevation, one at $+200^\circ$ and the other at -60 to $+20^\circ$ and $+200^\circ$. Highpass noise with high cutoffs (e.g., > 13 kHz) matched the strongly highpass characteristic of the $+200^\circ$ HRTF and matched, in the HRTFs from -60 to $+20^\circ$, the existence of energy at high frequencies and lack of energy in the mid frequencies.

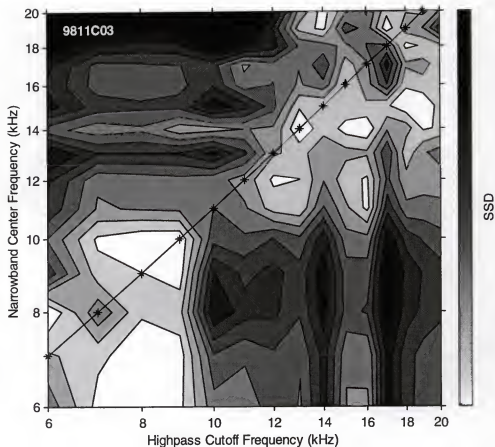


Figure 4.13. Sum of the squared differences (SSD) of network outputs. The contour plot represents the SSD between all pairs of distribution of network outputs for narrowband and highpass stimuli. Data of the distribution of network outputs are from the same unit (9811C03) shown in Figure 4.12. Highpass cutoff frequency is represented by the abscissa and narrowband center frequency is represented by the ordinate. White and light gray represent small SSD's and black and dark gray represent large SSD's. The line connected with asterisks (*—*) represents the frequencies at which the cutoff frequency of the highpass noise aligned with the lower cutoff of narrowband stimuli as in the upper panel of Figure 4.12.

In order to quantify the similarity of the network estimates of elevation for the spike patterns elicited by highpass and narrowband noise stimuli, we computed a sum of the squared differences (*SSD*) between all pairs of distribution of network outputs for both types of stimuli. A small *SSD* suggested similarity between the network outputs for the two types of stimuli. Figure 4.13 shows the *SSD*'s computed from the network outputs for the same unit (9811C03) shown in Figure 4.12. Lightness between black and white represents the *SSD* for each pair of the network estimates. Black and dark gray represent large *SSD*'s and white and light gray represent small *SSD*'s. The line connected with asterisks (*—*) represents the frequencies at which the cutoff frequency of the highpass noise aligned with the lower cutoff of narrowband stimuli as in the upper panel of Figure 4.12. That line fell in a region of minimum *SSD*'s.

We evaluated the hypothesis that network estimates of elevation based on highpass and narrowband noise are most similar when the low frequency cutoffs are equal. For each unit at each highpass cutoff frequency, we calculated the *SSD*'s between the network outputs for that highpass cutoff and every narrowband F_c . Next, we recorded the percentile rank of the *SSD* for the condition in which the highpass and narrowband lower cutoffs were equal. The null hypothesis predicts that the distribution of percentiles will be centered around 50%, whereas our hypothesis predicts that the distribution will lie considerably lower than 50%. Figure 4.14 plots the distribution of the percentile of matched *SSD* for 8 of the 15 highpass cutoff frequencies that we used. The distributions for the other 7 highpass noise cutoff frequencies are omitted for clarity but they were similar to those shown in Figure 4.14. Each panel represents the distribution, across all units recorded, for the highpass cutoff frequency that is indicated

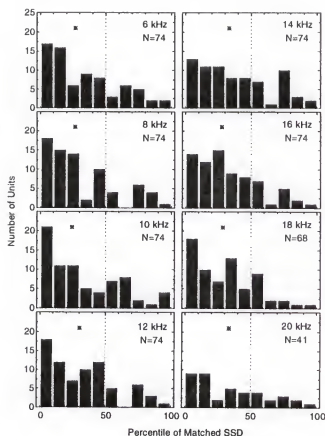


Figure 4.14. Distribution of percentile of matched SSD across the sample of units. Each panel represents data derived from one highpass cutoff frequency that is indicated in the upper right corner. For each unit at each highpass cutoff frequency, we calculated the SSD's between the network outputs for that highpass cutoff and every narrowband center frequency. The percentile of matched SSD was the percentile rank of the SSD for the condition in which the highpass and narrowband lower cutoffs were equal. The asterisk represents the median value of each distribution. The dashed line represents the chance-performance percentile of 50%.

in the upper-right corner of the panel. The asterisk represents the median value of each distribution. For all the 15 highpass noise cutoff frequencies, the median values of the percentile of matched SSD ranged from 20.0 to 38.2%. For all highpass noise cutoffs, 73.7% of our units had a percentile of matched SSD smaller than the chance-performance percentile of 50%. This result agrees with the result from human psychophysics (Macpherson and Middlebrooks 1999) that highpass and narrowband stimuli that have a common low-frequency cutoff tend to be referred to the same elevation.

Elevation Sensitivity by Spike Counts

In our previous reports, we showed that coding of sound-source azimuth and elevation by spike patterns is more accurate than coding by spike counts alone (Middlebrooks et al. 1998; Xu et al. 1998). Data from the present study confirmed such observations. We used the neural network procedure to classify the spike counts alone according to broadband source elevations and to compare the network performance with that using full spike patterns (Figure 4.15). Figure 4.15 shows data from the 40-dB fixed-level condition for the population of 389 units. The vertical and horizontal dotted lines represent the median value (50.4°) of the network performance using full spike patterns. When we used that value as a criterion to judge the network performance using spike counts alone, less than 10% (38/389) of the population would be considered elevation sensitive. For a large number of units, the network performance using spike counts alone was close to chance performance (i.e., median error = 65°). In fact, for 63.0% (245/389) of the sample of units, median errors obtained with spike counts alone were larger than 60° , whereas only 12.6% (49/389) of the units produced median error >

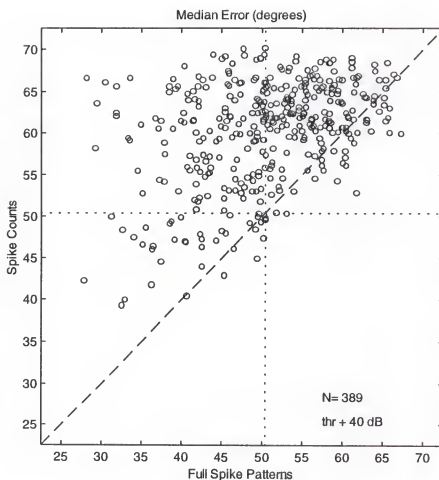


Figure 4.15. Accuracy of elevation coding by spike counts and by full spike patterns. Accuracy of coding was represented by the median error of the network outputs according to broadband sound-source elevation. Each symbol represents one A2 unit. Full spike patterns (abscissa) consisted of spike density functions expressed with 1-ms resolution. Spike counts (ordinate) were the total number of spikes in each density function. The dashed line on the main diagonal represents the equal performance line. The vertical and horizontal dotted lines represent the median values of the network performance with full spike patterns (50.4°).

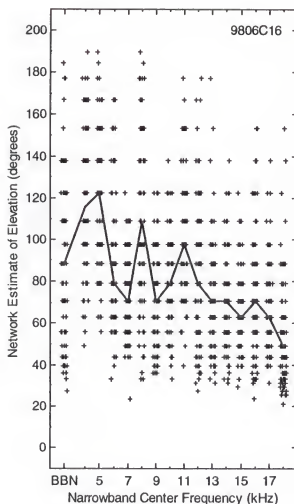


Figure 4.16. Network classification of spike counts elicited by narrowband sounds. The network analysis was based on spike counts elicited by narrowband sounds that varied in center frequency; the neural responses of the unit (9806C16) are shown in Figure 4.3. The neural network was trained with spike counts elicited by broadband noise presented from 14 elevations at 5 roving levels (20, 25, 30, 35, and 40 dB above threshold) and was tested with those elicited by narrowband noise at 30 dB above threshold. Each column of symbols represents network outputs for spike counts elicited by narrowband noise of a given center frequency as indicated along the abscissa. BBN indicates the network responses to spike counts elicited by broadband noise. All stimuli were presented from $+80^\circ$ elevation. The thick line indicates the median elevation of the network outputs for broadband noise and narrowband noise of various center frequencies.

60° with full spike patterns. Thus, our data indicated that information about sound-source elevation is to a large extent carried in the full spike patterns of cortical neurons.

Using spike counts alone as input to the neural networks, we evaluated the changes in elevation selectivity of unit response to narrowband stimuli. Figure 4.16 shows an example of the network estimates of elevation based on spike counts elicited by narrowband stimuli that varied in F_c ; the spike patterns and spike count tuning in response to narrowband stimulation of the unit (9806C16) is shown in Figure 4.3, B and E. The solid line in Figure 4.16 represents the median direction of the network outputs. In contrast to the network outputs based on full spike patterns (Figure 4.4), the network outputs based on spike counts showed very small variation with stimulus F_c and tended to scatter over a large range of locations. There was only a vague trend of change of the network-estimated elevations that followed the prediction by the localization model (background in Figure 4.4). In our sample of units, spike patterns consistently showed superior performance to spike counts in accounting for the accurate elevation coding of broadband sources and the systematic deviations under the condition of narrowband stimulation.

Discussion

The results confirm our previous observation that the spike patterns of units in area A2 can signal accurately the vertical locations of broadband sounds. The new finding of this study is that the spike patterns elicited by filtered stimuli, if interpreted as if they were the responses to broadband sounds, signal vertical locations that are systematically incorrect but that are predicted by an acoustic model. The computational

principles that lead to neuronal signals of correct and incorrect locations appear to correspond to the principles that underlie location judgments by human listeners. In this Discussion, we discuss the features of spectra that influence location judgements by human listeners and by cortical neurons, we evaluate the largely insignificant impact on elevation coding of notches in stimulus spectra, and we consider the importance of the magnitude and timing of neuronal responses for elevation coding.

Spectral Features and Elevation Coding

Human listeners would report that most if not all of the filtered sounds used in the present study sound different from broadband noise. Nevertheless, listeners appear to localize the filtered sounds as if they are broadband sounds that have been filtered by the listeners' own directional-dependent head-related transfer functions (HRTFs). In a study of narrowband localization, Middlebrooks (1992) found that the listeners exhibited systematic errors in elevation when asked to localize the narrowband sounds. A quantitative model based on the stimulus-HRTF correlation could successfully explain the systematic biases in the perception of elevation of narrowband sounds. The elevations of listeners' location judgments were those restricted regions in which the associated HRTFs correlated most closely with the stimulus spectra. Similar observations have been made in behavioral studies of cats. Huang and May (1996a) tested head orientation behavior in cats using 1/2-oct narrowband noise. They found, at least qualitatively, that cats oriented towards the spatial location where HRTF-filtering properties best matched the stimulus spectrum.

In the present study, we analyzed unit responses to filtered sounds as if they were responses to broadband sounds from particular locations. In that procedure, the neural

networks were trained with neural responses to broadband sounds from various elevations. We then used the trained neural networks to classify spike patterns elicited by various filtered noises and thereby to estimate the locations in elevation on the basis of match between the spike patterns elicited by filtered noise and broadband sounds. Our analysis procedure could be regarded as a physiological analogue of the behavioral procedure in which listeners localize filtered sounds.

The present study has demonstrated that the neuronal elevation selectivity is dependent on the center frequency of narrowband noise but independent of actual narrowband source location. These physiological data are consistent with psychophysical data from human listeners as well as from cats (human: Blauert 1969/1970; Hebrank and Wright 1974b; Middlebrooks 1992; Musicant and Butler 1985; cats: Huang and May, 1996a; Populin and Yin 1998). We adapted the localization model from previous human psychophysical studies (Middlebrooks 1992, 1999a) to predict the cats' localization judgments for narrowband sounds. The cortical neurons' spike patterns showed the same localization biases as behaving listeners in response to narrowband stimuli of various center frequencies. Therefore, the neurons' firing patterns might arise from a comparison between the stimulus spectra and a template of HRTFs. The cortical neurons that we studied might derive their elevation sensitivity from computational principles similar to those that underlie sound localization by human listeners.

The model of spectral shape recognition was most accurate in predicting neural responses to narrowband noise of mid-frequency F_c 's (i.e., 7 - 15 kHz) (Figure 4.9). The lower and higher frequency edges of the spectra of the 7- and 15-kHz narrowband noise (1/6-oct wide with 128-dB/oct slope) are 5.3 and 19.7 kHz (Figure 4.7, left panel). This

frequency range thus corresponded well to the mid-frequency range of 5–18 kHz that has been discussed as the most important frequency region for sound localization in cats (Rice et al. 1992; Neti et al. 1992; Huang and May 1996a). Rice and colleagues (1992) analyzed the HRTFs of cats and found that the mid-frequency region of 5–18 kHz contained spectral notches that varied systematically with sound-source elevation as well as azimuth. Neti and colleagues (1992) showed that an artificial neural network could be trained to perform the transformation from spectral information in HRTFs to a spatial map of sound-source locations. When bandlimited segments of frequency regions of the HRTFs were used as inputs to the neural network, they found that the mid-frequency region of 5–18 kHz provided the most robust localization cues. Recent behavioral studies in cat supported the importance of the mid-frequency spectra. Huang and May (1996a) reported that the cats could orient their heads to sound sources of mid-frequency bandpass noise of 5–18 kHz just as accurately as they did to broadband noise sources. Musicant and associates (1990) favored a slightly different mid-frequency range of 8–18 kHz as a spectral region that provided the most important spectral information for sound localization. Examining the HRTFs recorded from the eight cats that were used in the present study, we usually did not see significant variation of the spectral shape up to 6 or 7 kHz in the frontal locations. However, in the rear locations, spectral shape in the HRTFs started to vary at ~5 kHz (Figure 4.6). On the other hand, for most units, the spectral recognition model could not predict the neural responses to narrowband noise of F_c 's at low (4 and 5 kHz) or high frequencies (>17 kHz). Both low- and high-frequency regions of the HRTFs probably do not provide important spectral information for sound localization in the median plane. Our sample of units in area A2 usually did not respond

well to low-frequency sounds, as we reported previously (Xu et al. 1998). Consistent with other reports (Musicant et al. 1990; Rice et al. 1992), the high-frequency region (>17 kHz) in the HRTFs was highly complex and irregular. Although Huang and May (1996b) found that high frequency information might be used for minimal-audible-angle discrimination in the median plane by cats, such a frequency information apparently is not essential for vertical localization.

The model of spectral recognition performs spectral match between HRTFs and stimulus spectra (Middlebrooks 1992, 1999a). It does not reveal the most salient aspects of the spectra that are important for sound localization. Responses to narrowband noise might be based on increased energy at the center frequency or on slopes of the filter. The use of highpass noise in the present study provided us insights into the spectral cue processing of cortical neurons. Highpass and narrowband stimuli differs from each other in that they have very different spectral contents. They are similar in that they can share a common low cutoff frequency and positive slope.

We showed that the neural response patterns to highpass noise and narrowband noise resemble each other (Figures 4.11 to 4.14). This result suggests that the neurons' elevation selectivity is probably not based on the increased energy at the center frequency of narrowband noise but rather on the positive slopes in the spectra of both stimuli. Modeling studies of humans HRTFs demonstrated that the slopes of the HRTF spectra might provide more robust cues for sound localization than the spectra themselves (Macpherson 1998; Zakarauskas and Cynader 1993). A recent human psychophysical study in this laboratory provided evidence that human listeners tended to make equivalent localization judgments for narrowband and highpass sounds when the positive slopes in

the spectra of both stimuli match each other (Macpherson and Middlebrooks 1999). Therefore, both our electrophysiological and psychophysical findings indicate that the positive slopes in the spectra are probably a salient aspect of the spectral information that the HRTFs provide for vertical localization.

Influences of Spectral Notches on Elevation Coding

One of the prominent features in the HRTFs is the spectral notches in the mid-frequency region. In cat, the F_c 's of the spectral notches increase as the broadband noise source elevation increases in both frontal and rear locations (Figure 4.3). Detailed observations in this regard were made by different laboratories (Musicant et al. 1990; Rice et al. 1992). Psychophysical studies in human have show that elevation judgments could be influenced by bandstop filtering of white noise (Hebrank and Wright 1974b). Bloom (1977) also attempted to demonstrate that source elevation illusions in human could be created by notch filtering otherwise broadband signals. The notched noise was always presented at +60° elevation. When the F_c 's of the notched noise were varied from about 6 to 12 kHz, his listeners matched sound direction with flat spectrum sources placed between -45 and +40° in elevation. The F_c 's of the electronically-added notches corresponded to the frequency minima in the HRTFs of the phantom elevation. Under more natural localization conditions, however, narrow spectral notches generally produce illusions in elevation that are weak, at best (Macpherson 1998). No consistent evidence exists on whether cats' location judgments are influenced by notched noise.

In the present study, the responses of the A2 cortical neurons to notched stimuli appeared to be less sensitive to F_c than were responses to narrowband noise (Figure 4.3). Neural network analysis revealed that the spike patterns were more or less associated

with the actual location from which the notches were delivered (Figure 4.10).

Nonetheless, some variations in the network outputs were seen for certain notch F_c 's.

The variation in the network outputs, however, did not follow the prediction made from matching the notch F_c 's with the notch frequencies in the HRTFs. The model of spectral recognition that we proposed for the narrowband localization also failed to agree with the network outputs for the notch data. One possibility for these discrepancies is that the notch stimuli that we used (see METHODS from description) are physically different from the spectral notches that are present in the HRTFs. Another possibility is that a notch stimulus also contains flat spectral portions on either side of the notch and those flat spectral components might interact with the external-ear transfer function and thereby produce valid localization information to the brain. Therefore, at this stage, it still remains an open question whether a single notch (in the absence of other spectral cues) signals source elevation.

Elevation Coding by Spike Counts and Spike Timing

We have shown that elevation coding based on spike patterns that incorporate both spike counts and spike timing is more accurate than that based on spike counts alone (Figure 4.15, see also Xu et al. 1998). In fact, for most units, estimation of sound-source elevation using spike counts alone falls to near-chance performance level. We have also shown that, under conditions of narrowband stimulation, elevations signaled by spike patterns systematically follow the prediction of a localization model (Figure 4.4) whereas elevations signaled by spike counts alone show only vague trend of systematical biases that follow the model prediction (Figure 4.16). These results indicate that the

timing of spikes is an important information-bearing feature of the neural signal in the auditory cortex.

The difference in elevation coding between spike counts and spike patterns is perhaps a quantitative one rather than a qualitative one. Richmond and Optican (1987, 1990) represented cortical spike patterns in response to two-dimensional visual spatial patterns as a sum of successively more complex waveforms (principal components). It was shown that the first component, which was highly correlated with spike counts, carried about half of the information about the stimulus that was available in the spike patterns. Higher principal components, which represented spike timing, carried the other half of the total information. Our preliminary analysis of information-bearing elements along the same vein also showed that the first principal component accounted for about half of the variance across the spike patterns elicited by sounds presented from 360° of azimuth (Middlebrooks and Xu 1996). Nicolelis and colleagues (1998) recently found that the discrimination capability of area SII neural ensembles was significantly decreased when spike timing information was removed from the neuronal firing data. However, the discrimination capability using spike count alone was still above chance-performance level. It is possible that spike counts and spike timing code different stimulus parameters. For example, Gawne and colleagues (1996) find that in visual cortical neurons, spike counts seem to code stimulus orientation, whereas spike latencies code stimulus contrast. Nonetheless, it appears to be a general finding in the sensory cortex that spike timing carries additional information about stimuli in addition to what is carried by the spike counts.

Concluding Remarks

The present study confirms our previous report that the cortical neurons in area A2 code the location in elevation of a broadband sound source fairly accurately in their firing patterns but not as nearly accurately in the spike counts alone. We further show that the spike patterns are changed in some stereotyped manner when the broadband sounds are bandpass or highpass filtered. The association of neural responses to narrowband stimulation with sound-source elevations is a function of narrowband center frequency but independent of the actual narrowband source location. The neural responses elicited by narrowband noise tend to concentrate in the regions of elevation at which the spectral differences are found to be small. This is analogous to the tendency of human listener to orient to particular elevations when presented with narrowband noise. Also consistent with psychophysical work in human, highpass and narrowband sounds produce similar spike patterns that are classified into similar locations when the positive slopes of the spectra of both stimuli are at the same frequencies. The correlation that we see between physiology and behavior provides some insights into the functional significance of the firing patterns of cortical neurons. We do not have direct evidence that the neurons we studied in area A2 have a direct role in driving localization behavior. Our recordings from cortical area AES and preliminary data from area A1 indicate that sensitivity of spike patterns to sound-source elevation is not restricted to area A2, although A2 neurons manifest marginally superior performance to other cortical areas, possibly due to their broader frequency tuning properties (Xu et al. 1998). However, our results do demonstrate that sensitivity to broadband source elevation of A2 neurons breaks down under conditions of narrowband or highpass stimulation, as

seen in cat and human listeners. It is therefore adequate to conclude that the neuronal elevation sensitivity derives from mechanisms that are qualitatively similar to those that underlie localization behavior.

CHAPTER 5

SUMMARY AND CONCLUSIONS

Localization in the vertical plane and front/back discrimination involve using spectral shape cues provided by the filtering characteristics of the external ears. Previous studies have demonstrated that the spike patterns of auditory cortical neurons carry information about sound-source location in azimuth. The question arises as to whether those units integrate the multiple acoustical cues that signal the location of a sound source, or whether they merely demonstrate sensitivity to a specific parameter that co-varies with sound-source azimuth, such as interaural level difference. The experiments described in Chapter 3 addressed that issue by testing the sensitivity of cortical neurons to sound locations in the median vertical plane, where interaural difference cues are negligible. Auditory unit responses were recorded from 14 α -chloralose-anesthetized cats. We studied 113 units in the anterior ectosylvian auditory area (area AES) and 82 units in auditory area A2. Broadband noise stimuli were presented in an anechoic room from 14 locations in the vertical midline in 20° steps, from 60° below the front horizon, up and over the head, to 20° below the rear horizon, as well as from 18 locations in the horizontal plane. The spike counts of most units showed fairly broad elevation tuning. Averaged spike patterns were formed from the unit responses by averaging across multiple samples of 8 trials. An artificial neural network was used to recognize the spike patterns, which contain both the number and timing of spikes, and thereby to estimate the locations of sound sources in elevation. For each unit, the median error of neural-

network estimates was used as a measure of the network performance. For all 195 units, the average of the median errors was $46.4 \pm 9.1^\circ$, compared to the expectation of 65° based on chance performance. To address the question of whether sensitivity to sound pressure level (SPL) alone might account for the modest sensitivity to elevation of neurons, we measured SPLs from the cat's ear canal and compared the neural elevation sensitivity with the acoustical data. In many instances, the artificial neural network discriminated stimulus elevations even when the free-field sound produced identical SPLs in the ear canal. Conversely, two stimuli at the same elevation could produce the same network estimate of elevation, even when we varied sound-source SPL over a 20-dB range. There was a significant correlation between the accuracy of network performance in azimuth and in elevation. Most units that localized well in elevation also localized well in azimuth. Because the principal acoustic cues for localization in elevation differ from those for localization in azimuth, that positive correlation suggests that individual cortical neurons can integrate multiple cues for sound-source location.

Human and feline listeners can localize broadband sound accurately, but they make systematic errors in locations in the vertical plane when certain filters are applied to the source spectra. In the experiments described in Chapter 4, we studied the sensitivity of cortical neurons to the vertical locations of broadband and filtered sound sources. Stimuli consisted of 80-ms burst of broadband noise and noise filtered by narrow bandpass (narrowband), narrow band reject (notch) or highpass filters. Stimuli were presented from loudspeakers at 14 locations in the median plane, as in the experiments described in Chapter 3. We recorded responses from 389 units in the auditory cortical area A2 of 8 anesthetized cats, using the multichannel recording probes. We trained an

artificial neural network to recognize the spike patterns elicited by broadband noise and, thereby, to identify the source elevations. Then, the trained neural network was used to classify the spike patterns elicited by various filtered noises. The notch filters had little effect on elevation-specific responses of units. In contrast, the unit responses to narrowband noise of a particular center frequency or highpass noise of a particular cutoff tended to be classified around a particular elevation, regardless of the actual source location. Narrowband or highpass noise that varied in frequency content produced responses that were classified to varying elevations. Highpass and narrowband noise that shared a common low-frequency cut-off tended to produce similar spike patterns and similar neural-network outputs. We adapted to the cat a quantitative model that predicts human localization judgements of narrowband noise. That model, which incorporated external-ear transfer functions of each individual cat, could successfully predict the region in elevation that was associated with each narrowband center frequency.

In sum, our results show that spike patterns (spike counts and spike timing) of cortical neurons signal vertical sound locations correctly or systematically incorrectly under stimulus conditions that produce correct or incorrect localization by cats and human. This suggests that the cortical neurons that we studied derive their elevation sensitivity from computational principles similar to those that underlie sound localization behavior.

REFERENCES

- Andersen, P., Knight, L., & Merzenich, M. M. (1980). The thalamocortical and corticothalamic connections of AI, AII, and the anterior auditory field (AFF) in the cat: Evidence for two largely segregated systems of connections. *J. Comp. Neurol.*, 194, 663-701.
- Barlow, H. B. (1953). Summation and inhibition in the frog's retina. *J. Physiol. (Lond.)*, 119, 69-88.
- Barlow, H. B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception*, 1, 371-394.
- Barone, P., Clarey, J. C., Irons, W. A., & Imig, T. J. (1996). Cortical synthesis of azimuth-sensitive single-unit responses with nonmonotonic level tuning: A thalamocortical comparison in the cat. *J. Neurophysiol.*, 75(3), 1206-1220.
- Batteau, D. W. (1967). The role of the pinna in human localization. *Proc. Roy. Soc. Lond. B.*, 168, 158-180.
- Blauert, J. (1969-1970). Sound localization in the median plane. *Acustica*, 22, 205-213.
- Bloom, P. J. (1977). Creating source elevation illusions by spectral manipulation. *J. Audio Eng. Soc.*, 25, 560-565.
- Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig: Barth.
- Brugge, J. F., Reale, R. A., & Hind, J. E. (1996). The structure of spatial receptive fields of neurons in primary auditory cortex of the cat. *J. Neurosci.*, 16(14), 4420-4437.
- Brugge, J. F., Reale, R. A., Hind, J. E., Chan, J. C. K., Musicant, A. D., & Poon, P. W. F. (1994). Simulation of free-field sound sources and its application to studies of cortical mechanisms of sound localization in the cat. *Hear. Res.*, 73, 67-84.
- Butler, R. A., & Helwig, C. C. (1983). The spatial attributes of stimulus frequency in the median sagittal plane and their role in sound localization. *Am. J. Otolaryngol.*, 4, 165-173.
- Clarey, J. C., Barone, P., & Imig, T. J. (1994). Functional organization of sound direction and sound pressure level in primary auditory cortex of the cat. *J. Neurophysiol.*, 72(5), 2383-2405.

- Clarey, J. C., & Irvine, D. R. F. (1986). Auditory response properties of neurons in the anterior ectosylvian sulcus of the cat. *Brain Res.*, 386, 12-19.
- Clarey, J. C., & Irvine, D. R. F. (1990a). The anterior ectosylvian auditory field in the cat: I. An electrophysiological study of its relationship to surrounding auditory cortical fields. *J. Comp. Neurol.*, 301, 289-303.
- Clarey, J. C., & Irvine, D. R. F. (1990b). The anterior ectosylvian auditory field in the cat: II. A horseradish peroxidase study of its thalamic and cortical connections. *J. Comp. Neurol.*, 301, 304-324.
- Drake, K. L., Wise, K. D., Farraye, J., Anderson, D. J., & BeMent, S. L. (1988). Performance of planar multisite microprobes in recording extracellular single-unit intracortical activity. *IEEE Trans. Biomed. Engin.*, BME-35, 719-732.
- Efron, B., & Tibshirani, R. (1991). Statistical data analysis in the computer age. *Science*, 253, 390-395.
- Eggermont, J. J. (1998). Is there a neural code? *Neurosci. Biobehav. Rev.*, 22, 355-370.
- Fisher, H. G., & Freedman, S. J. (1968). The role of the pinna in auditory localization. *J. Aud. Res.*, 8, 15-26.
- Gardner, M. B., & Gardner, R. S. (1973). Problem of localization in the median plane: effect of pinnae cavity occlusion. *J. Acoust. Soc. Am.*, 53, 400-408.
- Gawne, T. J., Kjaer, T. W., & Richmond, B. J. (1996). Latency: Another potential code for feature binding in striate cortex. *J. Neurophysiol.*, 76(2), 1356-1360.
- Golay, M. J. E. (1961). Complementary series. *I.R.E. Trans. Information Theory*, 7, 82-87.
- Greene, T. C. (1929). The ability to localize sound: a study of binaural hearing in patients with tumor of the brain. *Arch. Surg.*, 18, 1825-1841.
- Hebrank, J., & Wright, D. (1974a). Are two ears necessary for localization of sound sources on the median plane? *J. Acoust. Soc. Am.*, 56, 935-938.
- Hebrank, J., & Wright, D. (1974b). Spectral cues used in the localization of sound sources on the median plane. *J. Acoust. Soc. Am.*, 56, 1829-1834.
- Henning, P., Tian, B., & Rauschecker, J. P. (1995). Piecewise continuous representation of azimuth and elevation in cat auditory cortex. *Abstr. Assoc. Res. Otolaryngol.* 18, 131.
- Hofman, P. M., Van Riswick, J. G. A., & Van Opstal, J. A. (1998). Relearning sound localization with new ears. *Nature Neurosci.*, 1(5), 417-421.

Huang, A. Y., & May, B. J. (1996a). Spectral cues for sound localization in cats: Effects of frequency domain on minimal audible angles in the median and horizontal planes. *J. Acoust. Soc. Am.*, 100(4), 2341-2348.

Huang, A. Y., & May, B. (1996b). Sound orientation behavior in cats. II. Mid-frequency spectral cues for sound localization. *J. Acoust. Soc. Am.*, 100(2), 1070-1080.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, 160, 106-154.

Humanski, R. A., & Butler, R. A. (1988). The contribution of the near and far ear toward localization of sound in the sagittal plane. *J. Acoust. Soc. Am.*, 83, 2300-2310.

Imig, T. J., Irons, W. A., & Samson, F. R. (1990). Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high-frequency primary auditory cortex. *J. Neurophysiol.*, 63, 1448-1466.

Imig, T. J., Poirier, P., Irons, W. A., & Samson, F. K. (1997). Monaural spectral contrast mechanism for neural sensitivity to sound direction in the medial geniculate body of the cat. *J. Neurophysiol.*, 78, 2754-2771.

Imig, T. J., & Reale, R. A. (1980). Patterns of cortico-cortical connections related to tonotopic maps in cat auditory cortex. *J. Comp. Neurol.*, 192, 293-332.

Jay, M. F., & Sparks, D. L. (1984). Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature*, 309, 345-347.

Jenkins, W. M., & Masterton, R. B. (1982). Sound localization: Effects of unilateral lesions in central auditory system. *J. Neurophysiol.*, 47, 987-1016.

Kistler, D. J., & Wightman, F. L. (1992). A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *J. Acoust. Soc. Am.*, 91, 1637-1647.

Klingon, G. H., & Bontecou, D. C. (1966). Localization in auditory space. *Neurol.*, 16, 879-886.

Knight, P. L. (1977). Representation of the cochlea within the anterior auditory field (AAF) of the cat. *Brain Res.*, 130, 447-467.

Knudsen, E. I. (1982). Auditory and visual maps of space in the optic tectum of the owl. *J. Neurosci.*, 2, 1177-1194.

Korte, M., & Rauschecker, J. P. (1993). Auditory spatial tuning of cortical neurons is sharpened in cats with early blindness. *J. Neurophysiol.*, 70, 1717-1721.

- Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., & Pitts, W. H. (1959). What the frog's eye tells the frog's brain. *Proc. I.R.E.*, 47, 1940-1951.
- Macpherson, E., & Middlebrooks, J. C. (1999). Sound localization illusions produced by source spectrum discontinuities. *Abstr. ARO Midwinter Meeting*, 22, 28.
- Macpherson, E. A. (1998). *Spectral cue processing in the auditory localization of sounds with wideband non-flat spectra*. Ph.D. dissertation, University of Wisconsin - Madison, WI.
- Mainen, Z. F., & Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268, 1503-1506.
- Makous, J. C., & Middlebrooks, J. C. (1990). Two-dimensional sound localization by human listeners. *J. Acoust. Soc. Am.*, 87, 2188-2200.
- May, B. J., & Huang, A. Y. (1996). Sound Orientation behavior in cats. I. Localization of broadband noise. *J. Acoust. Soc. Am.*, 100(2), 1059-1069.
- McClurkin, J. W., Gawne, T. J., Richmond, B. J., Optican, L. M., & Robinson, D. L. (1991). Lateral Geniculate neurons in behaving primates. I. Responses to two-dimensional stimuli. *J. Neurophysiol.*, 66(3), 777-793.
- Mehrgardt, S., & Mellert, V. (1977). Transformation characteristics of the external human ear. *J. Acoust. Soc. Am.*, 61, 1567-1576.
- Meredith, M. A., & Clemo, H. R. (1989). Auditory cortical projection from the anterior ectosylvian sulcus (field AES) to the superior colliculus in the cat: An anatomical and electrophysiological study. *J. Comp. Neurol.*, 289, 687-707.
- Merzenich, M. M., Knight, P. L., & Roth, G. L. (1973). Cochleotopic organization of primary auditory cortex in the cat. *Brain Res.*, 63, 343-346.
- Merzenich, M. M., Knight, P. L., & Roth, G. L. (1975). Representation of cochlea within primary auditory cortex in the cat. *J. Neurophysiol.*, 38, 231-249.
- Middlebrooks, J. C. (1992). Narrow-band sound localization related to external ear acoustics. *J. Acoust. Soc. Am.*, 92, 2607-2624.
- Middlebrooks, J. C. (1999a). Individual differences in external-ear transfer functions reduced by scaling in frequency. *J. Acoust. Soc. Am.*, in submission.
- Middlebrooks, J. C. (1999b). Virtual localization improved by scaling non-individualized External-Ear Transfer Functions in Frequency. *J. Acoust. Soc. Am.*, in submission.
- Middlebrooks, J. C., Clock, A. E., Xu, L., & Green, D. M. (1994). A panoramic code for sound location by cortical neurons. *Science*, 264, 842-844.

- Middlebrooks, J. C., Dykes, R. W., & Merzenich, M. M. (1980). Binaural response-specific bands in primary auditory cortex (AI) of the cat: Topographical organization orthogonal to isofrequency contours. *Brain Res.*, 181, 31-48.
- Middlebrooks, J. C., & Green, D. M. (1990). Directional dependence of interaural envelope delays. *J. Acoust. Soc. Am.*, 87, 2149-2162.
- Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Ann. Rev. Psychol.*, 42, 135-159.
- Middlebrooks, J. C., & Knudsen, E. I. (1984). A neural code for auditory space in the cat's superior colliculus. *J. Neurosci.*, 4, 2621-2634.
- Middlebrooks, J. C., Makous, J. C., & Green, D. M. (1989). Directional sensitivity of sound-pressure levels in the human ear canal. *J. Acoust. Soc. Am.*, 86, 89-108.
- Middlebrooks, J. C., & Pettigrew, J. D. (1981). Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to sound location. *J. Neurosci.*, 1, 107-120.
- Middlebrooks, J. C., & Xu, L. (1996). Information-bearing elements of spike trains in the cat's auditory cortex. *Soc. Neurosci. Abstr.*, 22, 1068.
- Middlebrooks, J. C., Xu, L., Eddins, A. C., & Green, D. M. (1998). Codes for sound-source location in nontopographic auditory cortex. *J. Neurophysiol.*, 80, 863-881.
- Middlebrooks, J. C., & Zook, J. M. (1983). Intrinsic organization of the cat's medial geniculate body identified by projections to binaural response-specific bands in the primary auditory cortex. *J. Neurosci.*, 3, 203-224.
- Miller, L. K., & Meredith, M. A. (1998). Field AES projections to auditory cortices. *Soc. Neurosci. Abstr.*, 24, 1880.
- Morel, A., & Imig, T. J. (1987). Thalamic projections to fields A, AI, P, and VP in the cat auditory cortex. *J. Comp. Neurol.*, 265, 119-144.
- Musicant, A. D., & Butler, R. A. (1985). Influence of monaural spectral cues on binaural localization. *J. Acoust. Soc. Am.*, 77, 202-208.
- Musicant, A. D., Chan, J. C. K., & Hind, J. E. (1990). Direction-dependent spectral properties of cat external ear: New data and cross-species comparisons. *J. Acoust. Soc. Am.*, 87, 757-781.
- Najafi, K., Wise, K. D., & Mochizuki, T. (1985). A high-yield IC-compatible multichannel recording array. *IEEE Trans. Electron. Devices*, ED-32, 1206-1211.

- Neti, C., Young, E. D., & Schneider, M. H. (1992). Neural network models of sound localization based on directional filtering by the pinna. *J. Acoust. Soc. Am.*, 92, 3140-3156.
- Nicolelis, M. A. L., Ghazanfar, A. A., Stambaugh, C. R., Oliveira, L. M. O., Laubach, M., Chapin, J. K., Nelson, R. J., & Kaas, J. H. (1998). Simultaneous encoding of tactile information by three primate cortical areas. *Nature Neurosci.*, 1, 621-630.
- Oldfield, S. R., & Parker, P. A. (1984). Acuity of sound localization: a topography of auditory space. II. Pinna cues absent. *Perception*, 13, 601-617.
- Oldfield, S. R., & Parker, P. A. (1986). Acuity of sound localisation: a topography of auditory space. III. Monaural hearing conditions. *Perception*, 15, 67-81.
- Palmer, A. R., & King, A. J. (1982). The representation of auditory space in the mammalian superior colliculus. *Nature*, 299, 248-249.
- Phillips, D. P., & Irvine, D. R. F. (1981). Responses of single neurons in physiologically defined primary auditory cortex (A1) of the cat: Frequency tuning and responses to intensity. *J. Neurophysiol.*, 45, 48-58.
- Phillips, D. P., & Irvine, D. R. F. (1982). Properties of single neurons in the anterior auditory field (AAF) of cat cerebral cortex. *Brain Res.* 248, 237-244.
- Populin, L. C., & Yin, T. C. (1998). Behavioral studies of sound localization in the cat. *J. Neurosci.*, 18, 2147-2160.
- Rajan, R., Aitkin, L. M., & Irvine, D. R. F. (1990a). Azimuthal sensitivity of neurons in primary auditory cortex of cats. II. Organization along frequency-band strips. *J. Neurophysiol.*, 64, 888-902.
- Rajan, R., Aitkin, L. M., Irvine, D. R. F., & McKay, J. (1990b). Azimuthal sensitivity of neurons in primary auditory cortex of cats. I. Types of sensitivity and the effects of variations in stimulus parameters. *J. Neurophysiol.*, 64, 872-887.
- Reale, R. A., & Imig, T. J. (1980). Tonotopic organization in auditory cortex of the cat. *J. Comp. Neurol.*, 192, 265-291.
- Reinoso-Suarez, F., & Roda, J. M. (1985). Topographical organization of the cortical afferent connections to the cortex of the anterior ectosylvian sulcus in the cat. *Exp. Brain Res.*, 59, 313-324.
- Rice, J. J., May, B. J., Spirou, G. A., & Young, E. D. (1992). Pinna-based spectral cues for sound localization in cat. *Hear. Res.*, 58, 132-152.

Richmond, B. J., & Optican, L. M. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. II. Quantification of response waveform. *J. Neurophysiol.*, 57(1), 147-161.

Richmond, B. J., & Optican, L. (1990). Temporal encoding of two-dimensional patterns by single units in primary visual cortex. II. Information transmission. *J. Neurophysiol.*, 64, 370-380.

Rieke, F., Warland, D., de Ruyter van Steveninck, R., & Bialek, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.

Roda, J. M., & Reinoso-Suarez, R. (1983). Topographical organization of the thalamic projections to the cortex of the anterior ectosylvian sulcus in the cat. *Exp. Brain Res.*, 49, 131-139.

Roffler, S. K., & Butler, R. A. (1968). Factors that influence the localization of sound in the vertical plane. *J. Acoust. Soc. Am.*, 43, 1255-1259.

Rose, J. E. (1949). The cellular structure of the auditory region of the cat. *J. Comp. Neurol.*, 91, 409-440.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In: D. E. Rumelhart, & J. McClelland, eds. *Parallel data processing*, 1, Chap. 8. Cambridge, MA: MIT Press: 318-620.

Sanchez-Longo, L. P., & Forster, F. M. (1958). Clinical significance of impairment of sound localization. *Neurol.*, 8, 119-125.

Schreiner, C. E., & Cynader, M. S. (1984). Basic functional organization of second auditory cortical field (AII) of the cat. *J. Neurophysiol.*, 51, 1284-1304.

Schreiner, C. E., & Mendelson, J. R. (1990). Functional topography of cat primary auditory cortex: Distribution of integrated excitation. *J. Neurophysiol.*, 64(5), 1442-1459.

Schreiner, C. E., & Sutter, M. L. (1992). Topography of excitatory bandwidth in cat primary auditory cortex: Single-neuron versus multiple-neuron recordings. *J. Neurophysiol.*, 68(5), 1487-1502.

Shadlen, M. N., & Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.*, 4, 569-579.

Shaw, E. A. G. (1974). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *J. Acoust. Soc. Am.*, 56, 1848-1861.

Slattery, W. H. I., & Middlebrooks, J. C. (1994). Monaural sound localization: Acute versus chronic impairment. *Hear. Res.*, 75, 38-46.

- Softky, W. R. (1995). Simple codes versus efficient codes. *Curr. Opin. Neurobiol.*, 5, 239-247.
- Sutter, M. L., & Schreiner, C. E. (1991). Physiology and topography of neurons with multi-peaked tuning curves in cat primary auditory cortex. *J. Neurophysiol.*, 65(5), 1207-1226.
- Sutter, M. L., & Schreiner, C. E. (1995). Topography of intensity tuning in cat primary auditory cortex: Single-neuron versus multiple-neuron recordings. *J. Neurophysiol.*, 73(1), 190-204.
- Victor, J. D., & Purpura, K. P. (1996). Nature and precision of temporal coding in visual cortex: A metric-space analysis. *J. Neurophysiol.*, 76(2), 1310-1326.
- Watkins, A. J. (1978). Psychoacoustical aspects of synthesized vertical locale cues. *J. Acoust. Soc. Am.*, 63, 1152-1165.
- Wightman, F. L., & Kistler, D. J. (1989). Headphone simulation of free field listening. I: Stimulus synthesis. *J. Acoust. Soc. Am.*, 85, 858-867.
- Wightman, F. L., & Kistler, D. J. (1997). Monaural sound localization revisited. *J. Acoust. Soc. Am.*, 101(2), 1050-1063.
- Winer, J. A. (1992). The functional architecture of the medial geniculate body and the primary auditory cortex. In: D. B. Webster, A. N. Popper, & R. R. Fay, eds. *The mammalian auditory pathway: Neuroanatomy*. New York: Springer-Verlag: 222-409
- Wise, L. Z., & Irvine, D. R. F. (1984). Interaural intensity difference sensitivity based on facilitatory binaural interaction in cat superior colliculus. *Hear. Res.*, 16, 181-187.
- Woodworth, R. S. (1938). *Experimental Psychology*. New York: Holt, Rinehart, and Winston.
- Wortis, S. B., & Pfeiffer, A. Z. (1948). Unilateral auditory-spatial agnosia. *J. Nerv. Ment. Dis.*, 108, 181-186.
- Xu, L., Furukawa, S., & Middlebrooks, J. C. (1998). Sensitivity to sound-source elevation in nontopographic auditory cortex. *J. Neurophysiol.*, 80, 882-894.
- Xu, L., & Middlebrooks, J. C. (1995). Coding of sound source elevation by firing patterns of auditory cortical neurons. *Soc. Neurosci. Abstr.*, 21, 667.
- Zakarauskas, P., & Cynader, M. S. (1993). A computational theory of spectral cue localization. *J. Acoust. Soc. Am.*, 94, 1323-1331.
- Zhou, B., Green, D. M., & Middlebrooks, J. C. (1992). Characterization of external ear impulse responses using Golay codes. *J. Acoust. Soc. Am.*, 92, 1169-1171.

BIOGRAPHICAL SKETCH

I was born in Changsha City, Hunan Province, China, in September 1963. In 1980, I began studying Medicine in Hengyang Medical College, Hengyang City, Hunan Province. In the fourth or fifth year of medical school, I decided to specialize in otolaryngology and to do research in inner ear diseases and in hearing science.

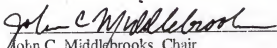
After I graduated from medical school in 1985, I was admitted to the graduate school of Capital University of Medical Sciences in Beijing. Under the supervision of Professors Yin-Shi Zhao, M.D., and Xiao-Lun Zhu, M.D., in the Department of Otolaryngology of Beijing Tongren Hospital and Beijing Institute of Otorhinolaryngology, I finished my thesis research on otoimmunology of Ménière's Disease. In 1988, I was appointed research assistant at Beijing Institute of Otorhinolaryngology. Meanwhile, I started my resident training in the Department of Otolaryngology, Beijing Tongren Hospital under the supervision of Professor Chan Liu, M.D., then the Director of Beijing Institute of Otorhinolaryngology.

In 1991, I was invited to study the problem of autoimmune inner ear diseases by Professor Carl R. Pfaltz, M.D., Chairman of the Department of Otorhinolaryngology of the University of Basel, Switzerland. For a period of a year and a half, I carried out a research project on the HLA-antigen linkage in patients with autoimmune inner ear diseases, in cooperation with Professor Wolfgang Arnold, M.D., from Luzent, Switzerland. I was also very fortunate to be able to work with Dr. Frances Harris,

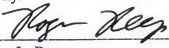
Ph.D., and Professor Rudolf Probst, M.D., the present Chairman of the Department, on otoacoustic emissions, a topic in which I had developed a new interest.

In fall 1992, I became a Ph.D. student at Dr. John Middlebrooks's laboratory at the Department of Neuroscience, University of Florida. The research topic was on the cortical neurophysiology of sound localization with special emphasis on the encoding of sound-source elevation by the spike patterns of the cortical neurons. In 1995, Dr. Middlebrooks accepted a new job at Kresge Hearing Research Institute, University of Michigan. I moved to Ann Arbor with him that summer and then finished the majority of my dissertation research there in the next three and a half years. Through my Ph.D. training with Dr. Middlebrooks, I have built a strong foundation for basic research in neuroscience. I would like to solidify such a foundation in the next few years and then carry on my own independent research in a direction that will be more clinical oriented and that will potentially benefit the health care of patients.

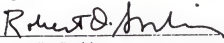
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


John C. Middlebrooks, Chair
Associate Professor of Neuroscience

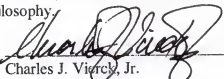
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Roger L. Reep
Associate Professor of Neuroscience

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Robert D. Sorkin
Professor of Psychology

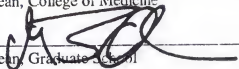
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Charles J. Viereck, Jr.
Professor of Neuroscience

This dissertation was submitted to the Graduate Faculty of the College of Medicine and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 1999


Dean, College of Medicine


Dean, Graduate School

UNIVERSITY OF FLORIDA



3 1262 08555 2916